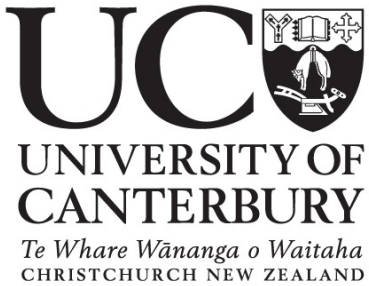
Studies into the allosteric regulation of α-isopropylmalate synthase

---

A thesis submitted in partial fulfilment of
the requirements for the degree
of
Doctor of Philosophy in Biochemistry
in the Department of Chemistry
by

Frances Helen Adam Huisman



August 2012

Abstract

α-Isopropylmalate synthase (α-IPMS) catalyses the first committed step in leucine biosynthesis in bacteria, including *Neisseria meningitidis* and *Mycobacterium tuberculosis*. It catalyses the condensation of α-ketoisovalerate (α-KIV) and acetyl coenzyme A (AcCoA) to form α-isopropylmalate (α-IPM). Like many key enzymes in biosynthesis, α-IPMS is inhibited by the end-product of the biosynthetic pathway, in this case leucine. α-IPMS is homodimeric, with monomers consisting of a $(\beta/\alpha)_8$-barrel catalytic domain, two subdomains and a C-terminal regulatory domain, responsible for binding leucine and providing feedback inhibition for leucine biosynthesis.

The exact mechanism of feedback inhibition in this enzyme is unknown, despite the elucidation of crystal structures with and without leucine bound. This thesis explores the nature of allosteric regulation in α-IPMS, including the effects of the regulatory domain and the importance of structural asymmetry on catalytic activity.

Chapter 2 details the characterisation of wild-type α-IPMS from *N. meningitidis* (*Nme*IPMS). This protein was successfully cloned, expressed and purified by metal-affinity and size-exclusion chromatography. *Nme*IPMS has similar characteristics to previously characterised α-IPMSs, being a dimer and demonstrating substrate binding affinities in the micromolar range. This enzyme has a turnover number of $13\,s^{-1}$ and is sensitive to mixed, non-competitive inhibition by the amino acid leucine. Small angle X-ray scattering experiments reveal that the solution-phase structure of this protein is likely similar to existing crystal structures of other α-IPMSs.

In Chapter 3, substitutions of residues potentially involved in the binding and transmission of the leucine regulatory mechanism are described. Most of these amino acid substituted variants reduce enzyme sensitivity to leucine, and one variant is almost entirely insensitive to this inhibitor. Another of these variants demonstrates an unexpected decrease in substrate affinity, despite the substituted residue being located

far from the active site.

The independence of α-IPMS domains is investigated in Chapter 4. The catalytic domains were isolated from *Nme*IPMS and the α-IPMS from *M. tuberculosis* (*Mtu*IPMS), and found to be unable to catalyse the condensation of substrates, despite maintaining the wild-type structural fold. Complementation studies with *Escherichia coli* cells lacking the gene for α-IPMS show that the truncated variants are unable to rescue growth in these cells. Binding of α-KIV in the truncated *Nme*-IPMS variant is much stronger than in the wild-type, and this may be the reason for lack of competent catalysis. A crystal structure was solved for the truncated variant of *Nme*IPMS and indicates that the regulatory domain is required for proper positioning of large regions of the protein. Two isolated regulatory domains from *Nme*IPMS were cloned, but with limited success in characterisation.

Finally, Chapter 5 describes substitutions made in *Mtu*IPMS to affect relative domain orientations within the protein. Dimer asymmetry is investigated by substituting residues at the domain interfaces. These substitutions did have some effect on catalysis and inhibition, but did not show any change in average solution-phase structure.

These results are drawn together in the greater context of allostery in general in Chapter 6, along with ideas for future research in this field. This chapter reviews the insights gained into protein structure from this thesis, particularly the importance of residues at protein domain interfaces. The asymmetry in the α-IPMS structure is discussed, along with small-molecule binding regulatory domains.

Acknowledgements

There are many people who must be thanked for their help and support during the creation of this thesis.

First and foremost, I would like to thank Emily Parker for her supervision and guidance. I am immensely grateful to her for always making time for me, no matter how many students she accumulates. Without her, I would never have had this experience.

I would like to acknowledge Heather Baker for her help and expertise, and Esther Bulloch for her generosity and enthusiasm. Thanks too, to Nayden Koon for giving me access to his previous work, and to Wayne Patrick for taking an interest in my research.

Considerable thanks go to Michael Hunter for discussing all things α-IPMS (and a great deal else) with me. I would also like to thank Tim Allison for, perhaps unwisely, always making himself available to answer questions.

Thank you to all the members, past and present, of the Parker Group for their support and friendship. Special thanks go to Tammie Cookson, Sarah Wilson-Coutts, Nicky Blackmore, Dmitri Joseph, Richard Hutton and Wanting Jiao for keeping me sane the last few years.

I would like to thank Sean Devenish for introducing me to the ways of the lab, and Grant Pearce and Andrew Muscroft-Taylor for their technical advice. Thanks also to all of our other colleagues in Biology for their friendly welcome whenever I find myself in their department.

I owe a lot to Ted Baker, Genevieve Evans and all the other members of the Maurice Wilkins Centre, for taking me in and making me feel at home when I was an earthquake refugee. I must particularly thank Christopher Squire for introducing me to his protein asymmetry project, and for providing financial support during my time in Auckland.

Much of my work could not have been done without the behind-the-scenes support of the excellent technical staff in the UC Chemistry and Biology Departments and at the Maurice Wilkins Centre. I am particularly thankful to Marie Squire for her technical expertise and to Martin Middleditch for giving up his time to help me find a rogue disulfide bond.

I would like to thank the UC College of Science for my doctoral scholarship, without which I would not have been able to undertake this degree. I am also grateful to the New Zealand Society of Biochemistry and Molecular Biology, the Biomolecular Interaction Centre and the Canterbury Branches of the New Zealand Federation of Graduate Women and the Royal Society of New Zealand for conference travel grants and other financial support.

Thank you to my family, who are always supportive of everything I do, even if they don't understand it.

Finally my greatest thanks, as always, go to Scott for all his support, encouragement and patience. I could not have done this without him.

Contents

List of Figures

List of Tables

List of Abbreviations

α-IPM α-isopropylmalate

α-IPMS α-isopropylmalate synthase

α-KB α-ketobutyrate

α-KIC α-ketoisocaproate

α-KIV α-ketoisovalerate

AcCoA acetyl coenzyme A

ACT aspartate kinase-chorismate mutase-tyrA (domain)

AHAIR acetohydroxyacid isomeroreductase

AHAS acetohydroxyacid synthase

ATP-PRTase adenosine triphosphate phosphoribosyltransferase

BCAT branched-chain aminotransferase

BCG Bacillus Calmette-Guérin

BTP 1,3-bis[tris(hydroxymethyl)methylamino]propane

cAMP cyclic adenosine monophosphate

CD circular dichroism

CMS citramalate synthase

CoA coenzyme A

DAH7PS 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase

DHAD dihydroxyacid dehydratase

DNA deoxyribonucleic acid

DSF differential scanning fluorimetry

DTP 4,4′-dithiodipyridine

DTT dithiothreitol

EDTA ethylenediaminetetraacetic acid

FRET fluorescence resonance energy transfer

GAPDH D-glyceraldehyde-3-phosphate dehydrogenase

HDX hydrogen/deuterium exchange

HEPES 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid

IMAC immobilised metal affinity chromatography

IPMD β-isopropylmalate dehydrogenase

IPMI α-isopropylmalate isomerase

IPTG isopropyl β-D-1-thiogalactopyranoside

ITC isothermal titration calorimetry

KNF Koshland-Némethy-Filmer

LB lysogeny broth

MAM methylthioalkylmalate synthase

MES 2-(*N*-morpholino)ethanesulfonic acid

MS mass spectrometry

***Mtu*IPMS** *Mycobacterium tuberculosis* α-IPMS

MWC Monod-Wyman-Changeux

NAD nicotinamide adenine dinucleotide

***Nme*IPMS** *Neisseria meningitidis* α-IPMS

NMR nuclear magnetic resonance

PCR polymerase chain reaction

PDB Protein Data Bank

PGDH phosphoglycerate dehydrogenase

PRCG Polak-Ribiere Conjugate Gradient

rms root mean squared

SAXS small angle X-ray scattering

SDS-PAGE sodium-dodecyl-sulfate polyacrylamide gel electrophoresis

SEC size-exclusion chromatography

TAE tris-acetate-ethylenediaminetetraacetic acid

TD threonine deaminase

TEV tobacco etch virus

ThDP thiamin diphosphate

TyrRS tyrosyl-tRNA synthetase

VNTR variable-number tandem repeat

WaterLOGSY water-ligand observed via gradient spectroscopy

Publications

Parts of this thesis have been published in the following publications:

Huisman, F. H. A., Hunter, M. F. C., Devenish, S. R., Gerrard, J. A., Parker, E. J., The C-terminal regulatory domain is required for catalysis by *Neisseria meningitidis* α-isopropylmalate synthase. *Biochemical and Biophysical Research Communications* **2010** 393(1):168–173

Huisman, F. H. A., Koon, N., Bulloch, E. M. M., Baker, H. M., Baker, E. N., Squire, C. J., Parker, E. J., Removal of the C-terminal regulatory domain of α-isopropylmalate synthase disrupts functional substrate binding. *Biochemistry* **2012** 51(11):2289–2297

Chapter 1

Introduction

1.1 Branched-chain amino acid biosynthesis

Antibiotics, by their nature, exploit the differences between pathogenic and mammalian cells. The ability to target cells of a particular organism or a particular class of organisms is what distinguishes antibiotics from poisons. Studying key biosynthetic pathways that are critical for life in bacteria, but are absent from mammals, allows for the development of targeted treatments for pathogenic infections that do not harm the infected host.

Among these key biosynthetic pathways are those responsible for production of amino acids — the building blocks of proteins and, by extension, cells. Organisms may generate these building blocks in one of two ways, either by the breakdown of ingested protein or by synthesis within the cell. Humans obtain ten of the twenty standard amino acids from diet; these ten are classified as the "essential" amino acids.[1] All three branched-chain amino acids — valine, leucine and isoleucine — are essential to humans but are synthesised in fungi,[2] plants[3] and bacteria.[4,5]

Several studies have shown that genetic knockouts of enzymes in the branched-chain amino acid biosynthetic pathway lead to auxotrophic mutants, unable to grow without amino acid supplementation (see Section 1.1.3). Thus, disruption of this pathway could be a potential treatment for pathogenic infection, making the branched-chain amino acid biosynthetic pathway a viable target for antibiotic design.

1.1.1 Valine and isoleucine biosynthesis

Pyruvate is the starting compound for the committed biosynthetic pathways of all three branched-chain amino acids, although the biosynthesis of isoleucine is also dependent on a threonine precursor in many organisms.[6] Valine and isoleucine are synthesised by near identical methods and share enzymes for many steps in their pathways (Figure 1.1). The dedicated leucine pathway branches off from α-keto-isovalerate (α-KIV), the immediate precursor to valine. Each of the first committed steps in these pathways are inhibited by one or more branched-chain amino acids.

The the first step towards isoleucine biosynthesis is the formation of α-ketobutyrate (α-KB) from threonine by the enzyme threonine deaminase (TD). TD is inhibited allosterically by isoleucine, and this inhibition can be reversed by valine.[7] An alternative, threonine-independent route to α-KB occurs in some bacteria via the citramalate pathway (shown in grey in Figure 1.1).[8] This pathway starts with the condensation of pyruvate and acetyl coenzyme A (AcCoA) by citramalate synthase (CMS) to form citramalate. Like TD, this enzyme is inhibited by isoleucine.[9] From citramalate a further two enzymes are required to form α-KB.

The next enzyme in branched-chain amino acid synthesis is acetohydroxyacid synthase (AHAS), which catalyses the reaction between α-KB and pyruvate to form α-aceto-α-hydroxybutyrate. This same enzyme can also catalyse the first committed step in valine biosynthesis, taking a second pyruvate molecule as a substrate in place of α-KB to form α-acetolactate.[10] AHAS is comprised of separate catalytic and regulatory subunits,[11,12] where association of the regulatory subunit is required for full activity in the catalytic site.[13] This enzyme is regulated by valine, isoleucine and leucine, often with a preference for valine and some synergy between inhibitors.[6,14] For example, AHAS from *Arabidopsis thaliana* is inhibited by each of isoleucine, valine and leucine, and also synergistically by combination of leucine with either of the other two branched-chain amino acids.[15]

From AHAS onward the biosynthetic pathways to isoleucine and valine run in parallel. First, α-acetolactate and α-aceto-α-hydroxybutyrate are isomerised and reduced by acetohydroxyacid isomeroreductase (AHAIR). The products of these reactions are then dehydrated by dihydroxyacid dehydratase (DHAD) and finally transaminated by branched-chain aminotransferase (BCAT) to form the amino acids. Although AHAS is the main point of regulation in the shared pathway, both AHAIR and DHAD may also be inhibited by valine and leucine in some organisms.[16]

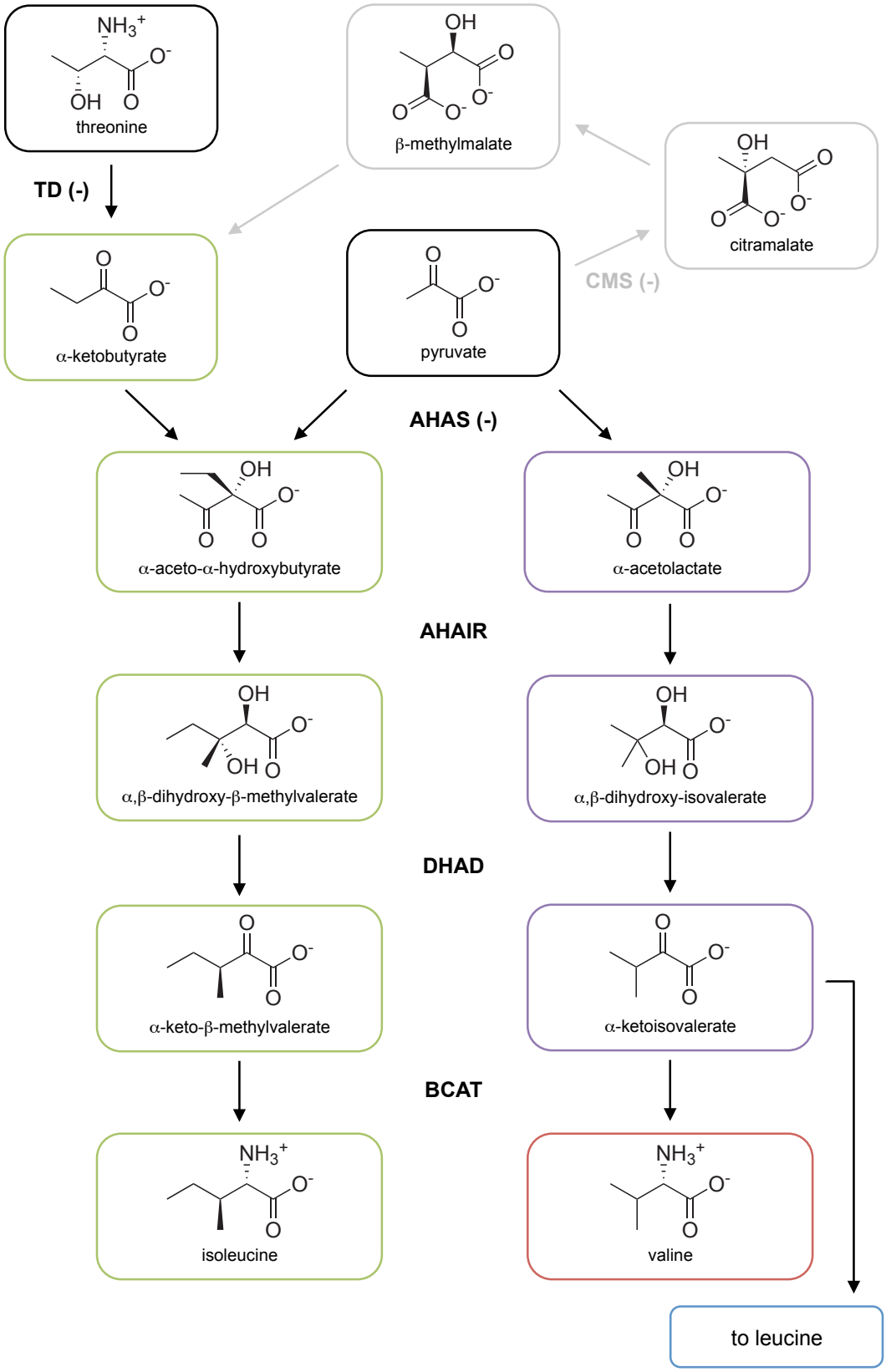


Figure 1.1: The branched-chain amino acid biosynthetic pathway. TD — threonine deaminase; CMS — citramalate synthase; AHAS — acetohydroxyacid synthase; AHAIR — acetohydroxyacid isomeroreductase; DHAD — dihydroxyacid dehydratase; BCAT — branched-chain aminotransferase. The alternative route to α-KB is shown in grey. The symbol (-) indicates points of regulation.

1.1.2 The leucine biosynthetic pathway

The last step in valine biosynthesis is the transamination of α-KIV by BCAT, but α-KIV is also the first substrate of the predominant bacterial leucine biosynthetic pathway.[17] This pathway is shown in Figure 1.2, with details of the chemistry involved.

The first committed step in leucine biosynthesis is the condensation of α-KIV and AcCoA to form α-isopropylmalate (α-IPM). This reaction is catalysed by α-isopropylmalate synthase (α-IPMS) and inhibited by leucine. From α-IPM, a hydroxyl group is transferred between adjacent carbons to form β-isopropylmalate, via an isopropylmaleate intermediate. This dehydration and rehydration is catalysed by α-isopropylmalate isomerase (IPMI). β-Isopropylmalate is then oxidised to α-isopropyl-β-oxosuccinate and decarboxylated to form α-ketoisocaproate (α-KIC), both steps catalysed by β-isopropylmalate dehydrogenase (IPMD). As for valine and isoleucine biosynthesis, the last step in the leucine pathway is transamination by BCAT to form the amino acid.

Some of the enzymes in leucine biosynthesis are also used in the threonine-independent formation of α-KB. The last three steps in this formation of α-KB from pyruvate are the dehydration of citramalate to form α-methylmaleate, followed by rehydration to β-methylmalate and oxidative decarboxylation to give the α-keto acid.[8,18,19] These reactions parallel those seen in the formation of α-KIC from α-IPM, and the same promiscuous enzymes (IPMI and IPMD) are used. The first reaction in this pathway is the condensation of pyruvate and AcCoA, which is catalysed by CMS and is analogous to the α-KIV/AcCoA condensation reaction of α-IPMS.

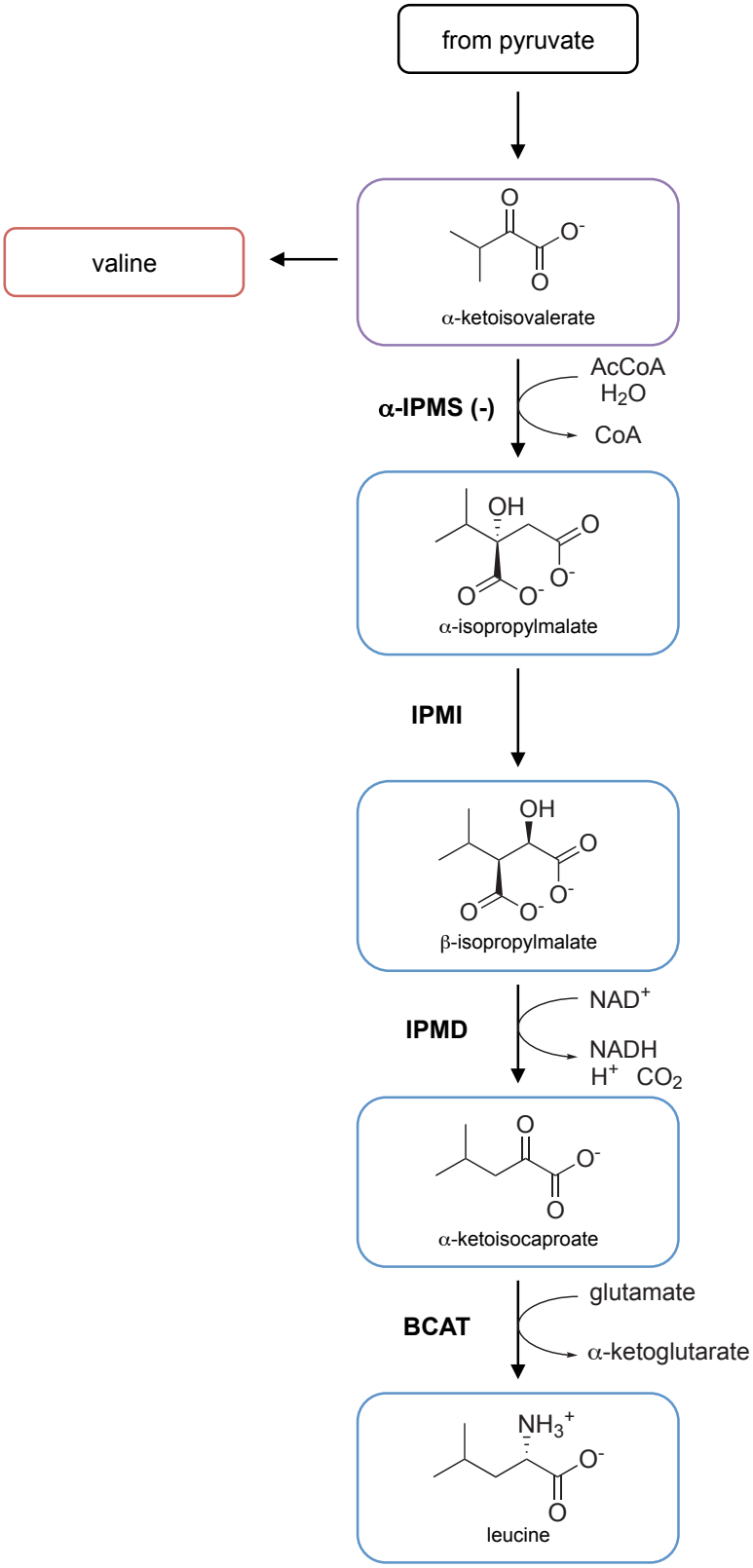


Figure 1.2: The leucine biosynthetic pathway. α-IPMS — α-isopropylmalate synthase; IPMI — α-isopropylmalate isomerase; IPMD — β-isopropylmalate dehydrogenase; BCAT — branched-chain aminotransferase. The symbol (-) indicates point of regulation.

1.1.3 Branched-chain amino acid biosynthesis as a drug target

Evidence from genetic mutation and inhibitor assays suggests that the branched-chain amino acid biosynthetic pathway is a viable drug target. Many of the enzymes in this pathway have proven critical to cell viability, with deletion or mutation of corresponding genes resulting in auxotrophy for one or more branched-chain amino acids (Table 1.1).

Enzyme (*gene*)	Auxotrophy	Species
TD (*ilvA*)	Ile	*Salmonella typhimurium*,[4] *Lemna minor*[3]
AHAS (*ilvB*)	Val, Ile, Leu	*Corynebacterium glutamicum*[20] *Mycobacterium tuberculosis*[5]
BCAT (*ilvE*)	Val, Ile, Leu	*Corynebacterium glutamicum*,[21] *Salmonella typhimurium*[4]
α-IPMS (*leuA*)	Leu	*Corynebacterium glutamicum*,[22] *Escherichia coli*,[23] *Lactococcus lactis*,[24] *Methanococcus maripaludis*,[25] *Salmonella typhimurium*,[2] *Neurospora crassa*[2]
IPMI (*leuC/D*)	Leu	*Mycobacterium bovis* BCG,[26,27] *Mycobacterium tuberculosis*[28]
IPMD (*leuB*)	Leu	*Bifidobacterium globosum*,[29] *Escherichia coli*,[23] *Lactococcus lactis*,[24] *Pseudomonas aeruginosa*[30]

Table 1.1: Examples of branched-chain amino acid auxotrophy, detailing genes affected, auxotrophy conferred and the species in which this was observed

Many enzymes in the branched-chain amino acid pathway are also known to be targets for herbicides. AHAS is a common target for commercially available compounds such as sulfonylaminocarbonyltriazolinones, triazolopyrimidines, pyrimidinylsalicylic acids, sulfonylureas and imidazolinones,[14] whereas AHAIR can be inhibited by a variety of potentially herbicidal compounds, including cyclopropane-1,1-dicarboxylate[31] and thiadiazoles.[32] Other inhibitors have also been demonstrated to affect TD, IPMI and IPMD.[14,33–35]

Despite its critical role in all three branched-chain amino acid pathways, BCAT is not an attractive inhibitor target due to its presence in mammals, where it is used in branched-chain amino acid catabolism.

Inhibition of leucine biosynthesis in the treatment of *M. tuberculosis*

The leucine biosynthetic pathway has been the target for development of several new vaccines against *M. tuberculosis*. Currently, vaccines are undertaken with *Mycobacterium bovis* Bacillus Calmette-Guérin (BCG), an avirulent strain of *M. bovis*, but this has disadvantages in that it may have pathogenic effects on individuals with compromised immune systems, and that inoculation with this strain gives false-positives in the tuberculin test for *M. tuberculosis* infection.[36]

Deletion of the gene *ilvB1* in *M. tuberculosis*, which encodes the major catalytic subunit of AHAS, has been shown to reduce cell growth of this pathogen in mice, but not affect cell viability of the bacteria in either mice or macrophages.[5] The persistence of these mutant *M. tuberculosis* cells within the host, coupled with defective growth, shows potential for vaccine development. Similarly, inactivation of gene *ilvE* in *M. tuberculosis*, encoding DHAD, reduces cell growth in mice and in culture.[37]

Several studies have found that the removal of IPMI activity from *M. bovis* BCG (through inactivation of the *leuD* gene) results in mutant cells unable to grow in mice or macrophages,[27] and cleared from the lungs and spleen in mice within seven weeks.[26] This strain protects guinea pigs from infection with *M. bovis* or *M. tuberculosis*, and does not give false-positive results for the tuberculin skin test in these animals.[36]

A double-knockout strain of *M. bovis* BCG lacking the *leuD* and *panCD* genes also protects guinea pigs from infection, with no adverse side-effects seen in simian immunodeficiency virus-infected Rhesus macaques.[38] *PanCD* is involved in the biosynthesis of pantothenate.

In addition to aiding the development of new vaccines, the branched-chain amino acid biosynthetic pathway has shown some promise as a target of antibiotic agents against *M. tuberculosis*. The AHAS inhibitor sulfometuron methyl is more effective than frontline anti-tuberculosis drugs against antibiotic-resistant *M. tuberculosis* strain T13704, lowering pathogenic cell numbers in the lungs, but not the spleen, of mice.[39] The AHAIR inhibitor *N*-isopropyloxalyl hydroxamate has also shown some antibiotic effects against antibiotic resistant *M. tuberculosis*. Both of these inhibitors were initially developed as herbicides, so it is possible that with further development they could lead to effective antibiotics.

1.2 α-Isopropylmalate synthase

α-IPMS (EC 2.3.3.13) catalyses the first committed step in leucine biosynthesis. It is the product of the gene *leuA* (*leu4* in *Saccharomyces cerevisiae*), and belongs to the Claisen-condensing family of enzymes, which catalyse the condensation of AcCoA with an α-keto acid. Other members of this enzyme family include malate synthase, methylthioalkylmalate synthase, *re*- and *si*-citrate synthases, homocitrate synthase and citramalate synthase. Many of these enzymes require a divalent metal ion for polarisation of the α-keto substrate, with the exception of *si*-citrate synthase, which uses a pair of histidine residues for this purpose.[40]

α-IPMSs from a range of organisms have been studied, including those from plants (notably *A. thaliana*[41]), fungi (*N. crassa*,[42] *S. cerevisiae*[43]) and bacteria (*M. tuberculosis*,[44–52] *S. typhimurium*,[53] *Corynebacterium glutamicum*[22]). A sequence alignment of some of the characterised α-IPMSs can be found in Appendix A. This alignment also indicates the major domains and secondary structure elements of the protein, as evidenced from crystal structures of α-IPMS from *M. tuberculosis* (*Mtu*IPMS).[51,54]

The most in-depth α-IPMS studies have been carried out on *Mtu*IPMS, due to the potential for drug development against this pathogen. The gene for *Mtu*IPMS is unusual amongst α-IPMSs in that it demonstrates sequence polymorphism, with different strains of *M. tuberculosis* containing different numbers of copies of a 57 base-pair tandem repeat. The number of copies of the repeat can be between two and 21, with two copies being the most common;[55] the deletion of this sequence does not affect enzyme activity or expression.[56] This variable-number tandem repeat (VNTR) is represented twice (residues 575–612) in the α-IPMS from *M. tuberculosis* strain H37Ra, which is the *Mtu*IPMS homologue described in this thesis.

1.2.1 Enzyme function

α-IPMS catalyses the condensation of α-KIV and AcCoA to form α-IPM (Figure 1.3). The enzyme provides maximum catalysis at slightly basic pH, with broad activity peaks around pH 8.5 in bacteria[41,53,57] and between pH 7.0 and 8.5 in fungi.[42,58] α-KIV and AcCoA have K_m values in the range of tens to hundreds of micromolar, with turnover numbers in single digits per second (Table 1.2). Studies into the effect of the VNTR in *Mtu*IPMS suggest that high numbers of repeats increase the binding affinity of the enzyme for both substrates while decreasing turnover number.[59] *Mtu*-

IPMS is also able to catalyse the α-KIV-independent hydrolysis of AcCoA with a K_m of 160 µM and a turnover number of 0.03 s^{-1}. The product CoA has been shown to inactivate α-IPMSs from *S. typhimurium* and *S. cerevisiae*.[60,61] Such inactivation is also seen in *S. cerevisiae* homocitrate synthase, though not in *S. cerevisiae* citrate synthase.[62]

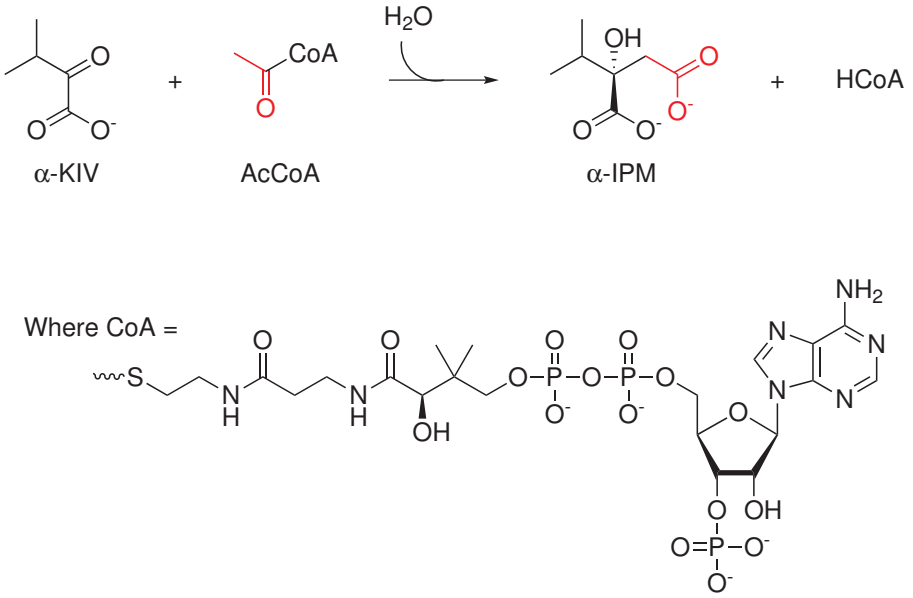


Figure 1.3: Reaction catalysed by α-IPMS

	K_m (µM)		
Organism	α-KIV	AcCoA	k_{cat} (s^{-1})
Mycobacterium tuberculosis[46]	12	136	3.5
Salmonella typhimurium[53]	60	200	
Saccharomyces cerevisiae[58]	16	9	
Neurospora crassa[42]		25	
Arabidopsis thaliana isozyme 1[41]	304	45	2.4
Arabidopsis thaliana isozyme 2[41]	279	16	2.3

Table 1.2: Kinetic data of characterised α-IPMSs. Some data not determined for some orthologues.

Metal dependence

α-IPMS requires a divalent metal ion for activity in all organisms studied. Mg^{2+} usually gives rise to highest catalytic activity in the enzyme, followed by Mn^{2+}.[41,43,47] Other divalent metal ions may activate or inhibit: Co^{2+} is activating for *M. tuberculosis* and *S. cerevisiae* α-IPMSs,[43,47] but has no effect on *A. thaliana* α-IPMS;[41] Ni^{2+} and Ca^{2+} are activating for *Mtu*IPMS but inhibitory for *A. thaliana* α-IPMS.[41,47]

These differences in metal dependency could be due to enzyme preference, but may also be due to methods of analysis. For instance, Zn^{2+} and Cd^{2+} were shown to be inhibitory for *Mtu*IPMS,[47] but later studies suggest that this was due to interference with the assay reporter compound.[63]

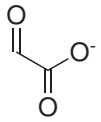
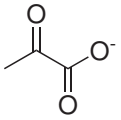
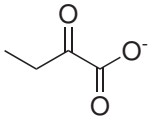
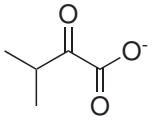
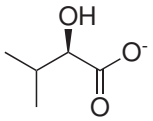
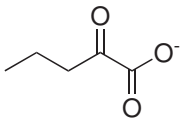
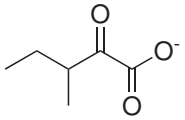
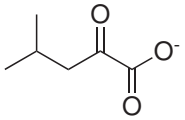
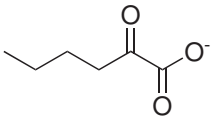
α-IPMSs from *S. cerevisiae*, *N. crassa* and *M. tuberculosis* show a requirement for monovalent cations, with a clear preference for K^+, although Na^+ and $NH_4{}^+$ are also activating.[42,47,58] These cations have no effect on substrate K_m values for *Mtu*IPMS, but do affect activation by Mg^{2+}. It is thought that monovalent cations may act to recruit divalent metal ions to the active site.[47]

Alternate substrates

In addition to α-KIV (the natural α-keto acid substrate) many α-IPMSs are also able to catalyse the condensation of AcCoA with pyruvate, α-KB or α-ketovalerate (Table 1.3). Many of these enzymes are inhibited by α-KIC, including those from *Alcaligenes eutrophus*,[64] *S. typhimurium*,[53] and *S. cerevisiae*.[58] α-KIC is the penultimate compound in the leucine biosynthetic pathway.

The two isozymes of α-IPMS from *A. thaliana* seem to have a greater tolerance for chain length variation than the bacterial and fungal enzymes, and can utilise glyoxalate, α-keto-β-methylvalerate, α-KIC and α-ketocaproate as weak substrates.[41]

α-Hydroxyisovalerate has been shown to inhibit *Mtu*IPMS, confirming the importance of the α-keto group in the α-keto acid substrate.[46] α-IPMS has a very strict tolerance for AcCoA, and generally will not accept other acyl-CoA analogues as alternate substrates. Some slight condensation activity has been observed with propionyl-CoA, however the majority of free CoA released by the enzyme in the presence of this substrate is due to α-KIV-independent hydrolysis. This is also true of the longer acyl-CoA analogue crotonyl-CoA.[46]

Compound	Structure	Activity [*see footnote*]
Glyoxalate		Weak substrate [f] No activity [b]
Pyruvate		Substrate [a–f]
α-Ketobutyrate		Substrate [a–f]
α-Ketoisovalerate		Natural substrate
(*S*)-α-Hydroxyisovalerate		Inhibitor [b]
α-Ketovalerate		Substrate [a–c,f] Inhibitor [f]
α-Keto-β-methylvalerate		Weak substrate [f] No activity [b]
α-Ketoisocaproate		Weak substrate [f] Inhibitor [a–d,f]
α-Ketocaproate		Weak substrate [f]

α-IPMSs from:

(a) *Alcaligenes eutrophus*[64]

(b) *Mycobacterium tuberculosis*[46]

(c) *Salmonella typhimurium*[53]

(d) *Neurospora crassa*[42]

(e) *Saccharomyces cerevisiae*[58]

(f) *Arabidopsis thaliana*[41]

Table 1.3: α-Keto acid specificity of α-IPMS

Reaction mechanism

The divalent metal ion in α-IPMS is proposed to coordinate the carbonyl groups of α-KIV and to polarise the α-keto group during catalysis, while it is thought that a monovalent cation may introduce structural changes to bind the divalent metal ion and the substrate more tightly.[47] However, not all variants of α-IPMS require monovalent cations, and hypotheses about any possible function and location are entirely conjecture at this point as there is no sign of such a cation in the solved crystal structures.[51]

The α-IPMS enzymatic reaction proceeds through enolisation of AcCoA, followed by condensation to form α-isopropylmalyl-CoA (Figure 1.4). The CoA moiety is then hydrolysed to give α-IPM. This reaction is proposed to involve one acidic and two basic amino acids within the α-IPMS active site.[46]

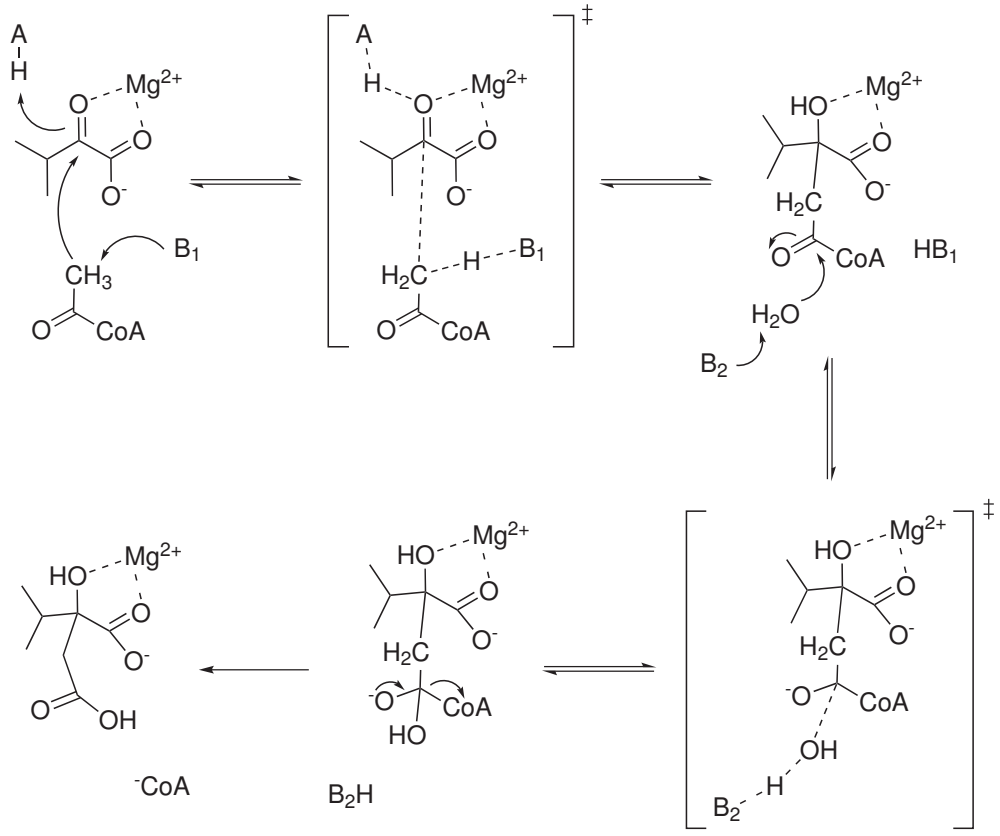


Figure 1.4: Detailed mechanism of α-IPMS reaction, as proposed by de Carvalho *et al.*[46] A, B_1 and B_2 are proposed acidic and basic residues in the enzyme active site.

Feedback inhibition by leucine

Like many enzymes which catalyse the first committed step in a biosynthetic pathway, α-IPMS demonstrates feedback inhibition by the pathway end-product, in this case leucine.[42]

Leucine has similar affinity for both the free and substrate-bound forms of *Mtu*IPMS, with a K_i of $8 \pm 1\,\mu M$ and a K_i' of $22 \pm 2\,\mu M$ with respect to α-KIV.[45] Inhibition is mixed and slow-onset in this enzyme, indicating a possible isomerisation event from the Enz·Leu complex to a more tightly bound Enz*·Leu complex. This complex has an inhibition constant (K_i*) of $3.6 \pm 2.2\,\mu M$, calculated from the linear, steady-state phase of leucine inhibition kinetics, as opposed to the initial velocity rates used to find the K_i and K_i'.

Inhibition in many α-IPMSs is pH dependent, with weaker inhibition evident at higher pH values.[42,53] The number of VNTRs is also important in *Mtu*IPMS regulation, as no inhibition is observed in *Mtu*IPMS isozymes with large numbers of these repeats.[59]

1.2.2 Overall structure

The only reported full-length α-IPMS crystal structures so far determined are from *M. tuberculosis*.[51] These structures show α-KIV, citrate, α-KIC or bromopyruvate bound at the active site (Protein Data Bank [PDB] codes 1SR9, 3HQ1, 3HPS and 3HPZ) or leucine bound in the regulatory domain (PDB code 3FIG). No unliganded or AcCoA-bound full-length structures have yet been solved.

*Mtu*IPMS is a 70 kDa monomer that forms a dimer in solution and in the crystal structure (Figure 1.5).[46,51] Other organisms have been shown to possess trimeric or tetrameric α-IPMSs.[41,42,65]

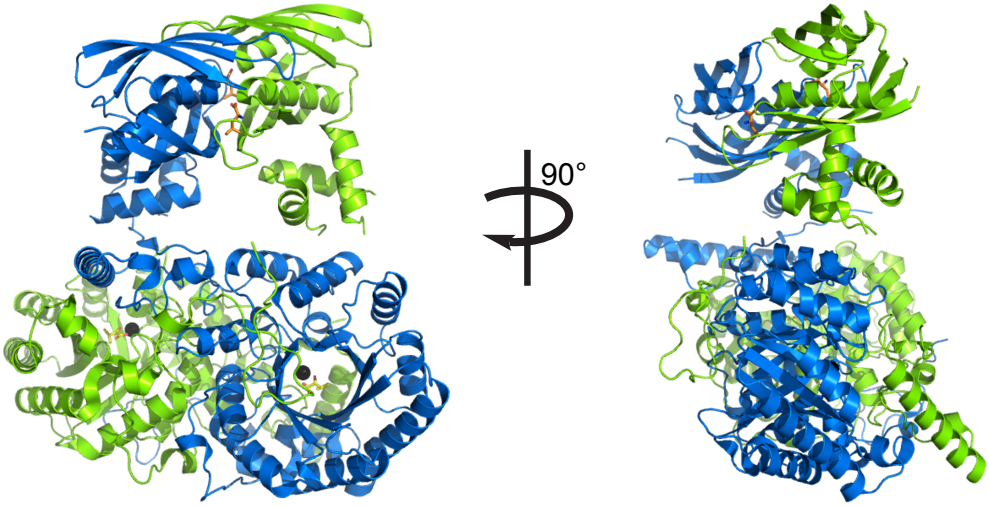
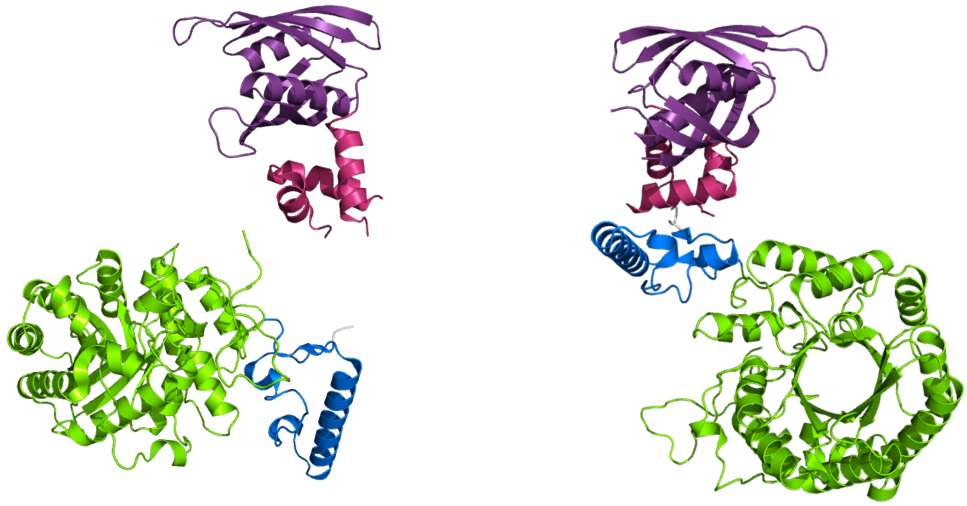


Figure 1.5: Crystal structure of α-IPMS from *M. tuberculosis* (PDB code 1SR9). Monomers are shown in different colours. α-KIV (yellow) and Zn^{2+} (black) are shown in the active sites. Leucine (orange) has been overlaid from structure 3FIG (leucine-bound *Mtu*IPMS).

Each monomer of *Mtu*IPMS is composed of two domains and two subdomains, separated by a flexible linker (Figure 1.6). The largest protein domain is the N-terminal, catalytic $(\beta/\alpha)_8$-barrel. This domain contains the active site, with substrate and divalent metal-ion binding sites at the C-terminal end of the barrel. Next is the small subdomain I, formed from two short β-strands and an α-helix, which crosses over in the dimer to sit over the active site of the adjacent monomer, and contributes two residues into the potential AcCoA binding cavity. The 9-residue flexible linker region is disordered in the crystal structures, and connects subdomain I to subdomain II. This second subdomain comprises three α-helices and is intimately associated with the C-terminal regulatory domain, itself composed of two $(\beta\beta\beta\alpha)_2$

units arranged so as to sandwich the α-helices between the β-sheets. The regulatory domain contains two leucine binding sites at its dimeric interface and two VNTRs that are undefined in the crystal structures.



(a) Monomer A **(b)** Monomer B

Figure 1.6: Domains in *Mtu*IPMS, shown in both conformations observed in the asymmetric dimer of crystal structure 1SR9. Catalytic domain in green, subdomain I in blue, subdomain II in pink and the regulatory domain in purple. The disordered 9-residue linker region connects the two subdomains.

Asymmetry within the protein dimer is evident in all full-length crystal structures of *Mtu*IPMS. This asymmetry is due to different relative positions in each monomer of the C-terminal domains (the regulatory domain and subdomain II) from the N-terminal domains (subdomain I and the catalytic domain), as shown in Figure 1.6. The structural difference appears to be mediated by the flexibility of the undefined linker region.

1.2.3 Active site

α-KIV and metal ions

A metal ion is essential for catalysis in α-IPMS, as it coordinates and polarises the carbonyl groups of α-KIV during the reaction. Crystal structures have elucidated the binding of the divalent metal ion and the substrate α-KIV (Figure 1.7), however the position of the monovalent cation required for activity in some α-IPMSs remains undefined.

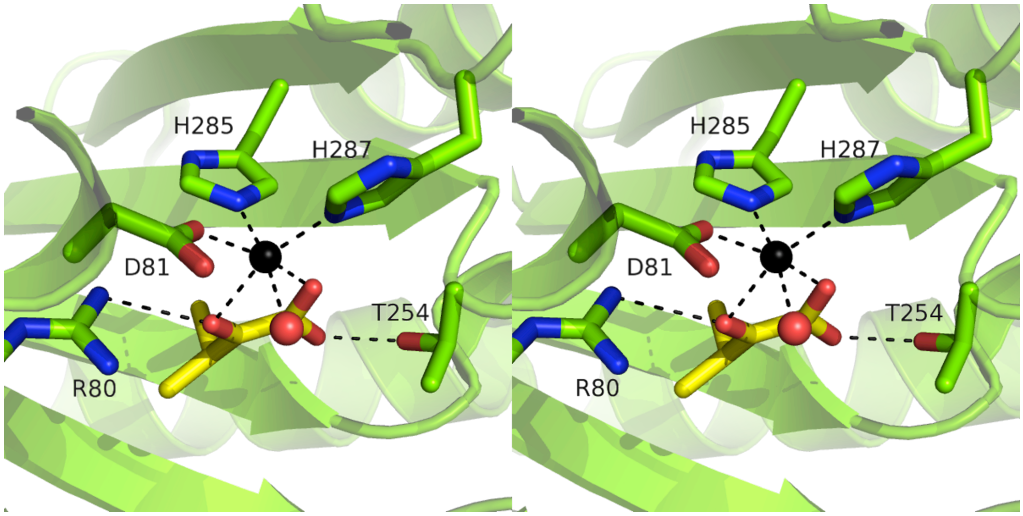


Figure 1.7: Stereo diagram of α-KIV binding in *Mtu*IPMS structure 1SR9. α-KIV is shown in yellow, Zn^{2+} in black and water in red.

In *Mtu*IPMS crystal structure 1SR9, a Zn^{2+} ion coordinates with two carbonyl groups of α-KIV, a water molecule and the side chains of Asp81, His285 and His287. α-KIV forms two other hydrogen bonds with Arg80 and Thr254. All of these metal- and substrate-binding residues are highly conserved in α-IPMS enzymes, except His287 which is a glutamine in both *A. thaliana* isozymes. The methyl groups of α-KIV are surrounded by Leu143, His167, Ser216, Asn250 and Pro252, which are also very well conserved and dictate the α-keto acid specificity of the enzyme. The equivalent histidine, serine and asparagine have been substituted for shorter amino acids (alanine or glycine) in *E. coli* α-IPMS to create a variant with increased promiscuity towards α-keto acids.[66] Increased promiscuity was also seen in a double substitution of the equivalent serine and proline (both to glycine) in *A. thaliana* α-IPMS.[67]

Binding of AcCoA

No crystal structure of α-IPMS has been solved with AcCoA bound, however such a structure does exist for the related enzyme CMS from *Leptospira interrogans*.[68] This enzyme catalyses the condensation of AcCoA and pyruvate as the first step in the threonine-independent synthesis of isoleucine, and is inhibited by isoleucine in a similar manner to leucine inhibition of α-IPMS.

The active site of CMS is very similar to that of *Mtu*IPMS (Figure 1.8), with a Zn^{2+} ion coordinating with two histidine residues, an aspartate, a water molecule and two oxygens of the α-keto acid substrate (in this case pyruvate). Pyruvate forms hydrogen bonds with a threonine and an arginine, just as α-KIV does in *Mtu*IPMS. A molecule of AcCoA sits in the cavity adjacent to pyruvate, where the acetyl oxygen forms a hydrogen bond to Arg16 while the methyl carbon lies between Glu146 and pyruvate. It is proposed that Glu146 acts as a catalytic base to enolise AcCoA in CMS, while Arg16 stabilises this enol form.[68] These two residues are well conserved and are equivalent to Glu218 and Arg80 in α-IPMS.

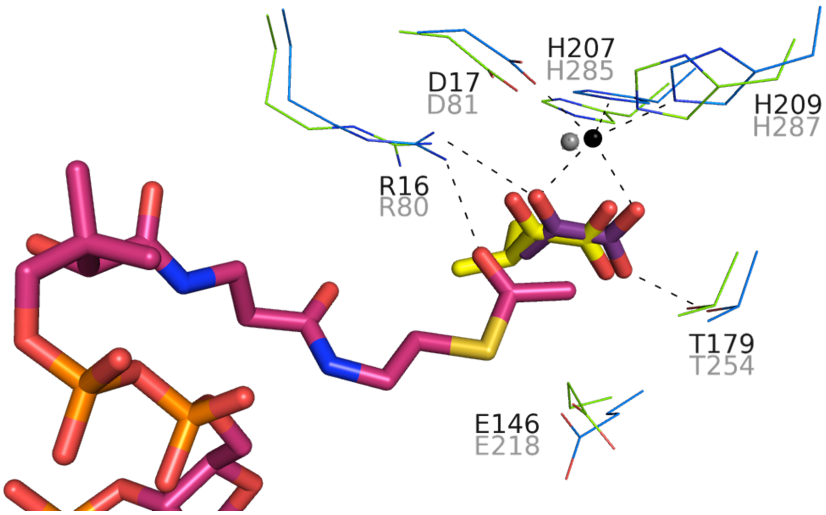


Figure 1.8: Active site of CMS (PDB code 3BLI — blue) overlaid with that of *Mtu*IPMS (1SR9 — green). AcCoA is shown in pink, pyruvate in purple and α-KIV in yellow. Zn^{2+} is shown in black (CMS) and grey (*Mtu*IPMS). Key residues are labelled for CMS (black) and α-IPMS (grey), and bonding in CMS is indicated.

Modelling studies with AcCoA in *Mtu*IPMS have indicated that the acetyl moiety binds in much the same way as in CMS. The acetyl oxygen forms a hydrogen bond with Arg80, while the methyl carbon sits between invariant residue E218, α-KIV and the subdomain I residue His379′ (Figure 1.9),[51] where the prime symbol (′) denotes

residues that cross over from the adjacent monomer. The methyl carbon and the carbonyl oxygen of the modelled AcCoA displace water molecules observed in the crystal structure.

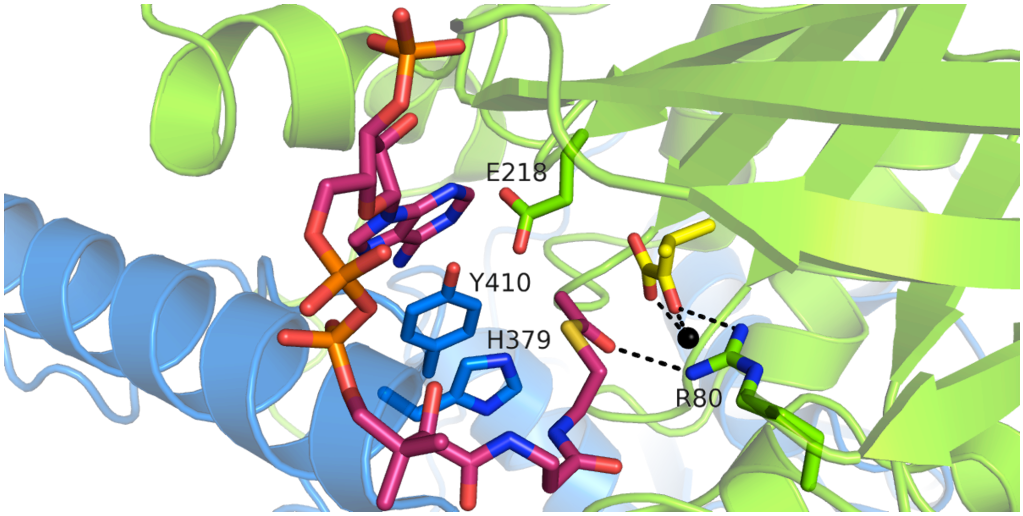


Figure 1.9: Modelled AcCoA binding in *Mtu*IPMS. Each monomer is shown in a different colour. AcCoA is shown in pink, α-KIV in yellow and Zn^{2+} in black.

Two basic residues and one acidic residue are necessary for catalysis in *Mtu*IPMS,[46] but the nature of these residues is unknown due to the number of conserved ionisable active site residues, including Arg80, Asp81, Glu218, Glu317, Arg318, His379′ and Tyr410′.

The modelled structure suggests that enolisation of the acetyl carbon could be achieved by Glu218 or His379′, while Arg80 acts to stabilise the acyl carbonyl group during enolisation. Analogy with CMS would suggest that Glu218 is the catalytic base in *Mtu*IPMS, and substitution of Glu218 with alanine reduces the enzyme k_{cat} 40-fold.[48] Similar substitution of His379′ for alanine has a far less pronounced effect on catalysis, however the fact that neither of these substitutions completely abolish activity in *Mtu*IPMS suggests either that there is some redundancy in the residue responsible for enolisation, or that this catalysis is promoted by one of the other highly conserved residues in the AcCoA binding pocket such as Glu317, Arg318 or Tyr410′. Interestingly, the E218A substitution also increases the leucine affinity of the free enzyme 10-fold, suggesting that the substitution of this residue may affect the distant leucine-binding site. Further substitutions of Arg80, Arg318 and Tyr410′ to alanine in *Mtu*IPMS have resulted in loss of catalytic activity.[54]

1.2.4 Structural elements of regulation

The leucine binding site

Leucine binding in *Mtu*IPMS has been elucidated by the crystal structure 3FIG.[51] Two leucines bind at the dimer regulatory-domain interface, as shown in Figure 1.10. The carboxyl group of leucine forms hydrogen bonds with the main-chain functional groups of Ala536 from one monomer and Ile627′ from the other. The leucine amino group forms bonds with the side-chain carbonyl of Asn532 and the main-chain carbonyl of Ala565′. The hydrophobic end of the inhibitor is surrounded by Val551, Tyr554, and Leu535.

Isothermal titration calorimetry (ITC) shows that leucine binds to *Mtu*IPMS with a K_D of 3.0 ± 0.6 µM,[48] comparing favourably with the K_i* of 3.6 ± 2.2 µM found for steady-state inhibition kinetics.[45]

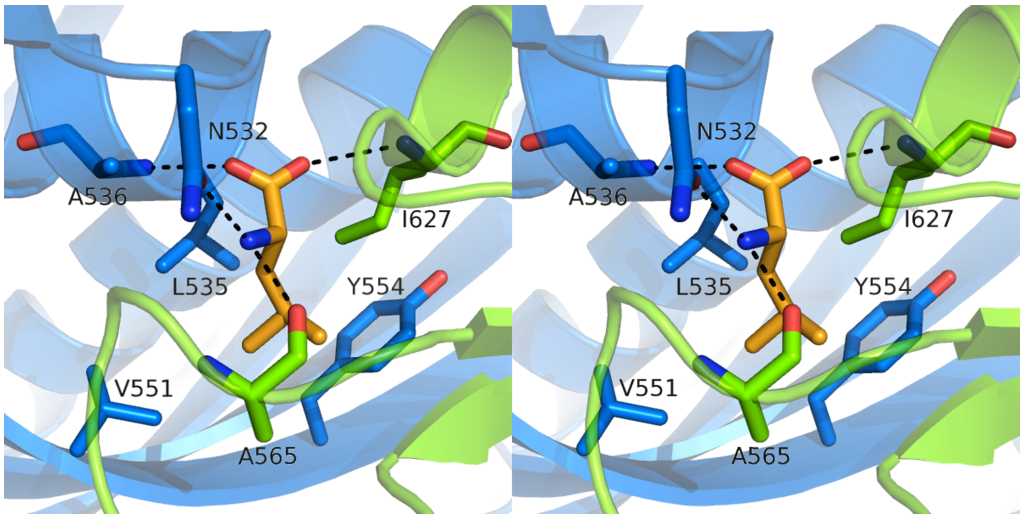


Figure 1.10: Stereo diagram of leucine binding in *Mtu*IPMS from crystal structure 3FIG.[51] Each monomer is shown in a different colour, leucine is shown in orange.

There are several conserved residues in the regulatory domain, most notably a G-x-G-P-[VIL] motif (residues 531–535 in *Mtu*IPMS) that forms a loop in the leucine binding site. The variable residue in this motif is an asparagine in *Mtu*IPMS and contributes a hydrogen bond to the leucine amino group, while the leucine residue forms part of the hydrophobic binding pocket around the inhibitor. Substitution of either of the glycines in this motif results in leucine insensitivity in *S. cerevisiae* α-IPMS.[69] Similar results are observed for the motif proline in *E. coli* α-IPMS.[70]

Most hydrogen bonds to leucine in *Mtu*IPMS occur through protein main-chain functional groups, from residues that are for the most part poorly conserved. The exception to this is Ile627, which is conserved as isoleucine or valine, and forms part of the hydrophobic pocket in addition to coordinating with the leucine carboxyl moiety through its main-chain amino group. Other conserved residues in the hydrophobic pocket are an invariant tyrosine (Tyr554 in *Mtu*IPMS) and two conserved branched-chain amino acids (Val551 and Val571 in *Mtu*IPMS).

Conserved residues in the wider *Mtu*IPMS regulatory domain include Ala541 and Ala634, both in the domain α-helices, and Ser631, which forms a stabilising hydrogen bond to Ile627 in the inhibitor binding pocket. Two invariant glycines (Gly620 and Gly622) are found at the interface of the regulatory domain with subdomain II, while a third (Gly562) is located in a flexible loop covering the binding site.

The leucine binding site flexible loop is formed by residues Met559–Gln566, and has recently been found to be the reason for the slow-onset nature of leucine inhibition observed in *Mtu*IPMS.[44] A slight change in position of the loop is the only structural change between leucine-bound and unbound crystal structures, and alanine-scanning mutagenesis suggests that the two-stage inhibition kinetics in this enzyme can be attributed to closing of these residues over the inhibitor binding site. Deletion of a residue in this loop has conferred leucine resistance in *S. cerevisiae* α-IPMS, as does substitution of an alanine (Ala552 in *S. cerevisiae* α-IPMS) for a polar threonine.[69] This alanine is equivalent to Ala568 in *Mtu*IPMS.

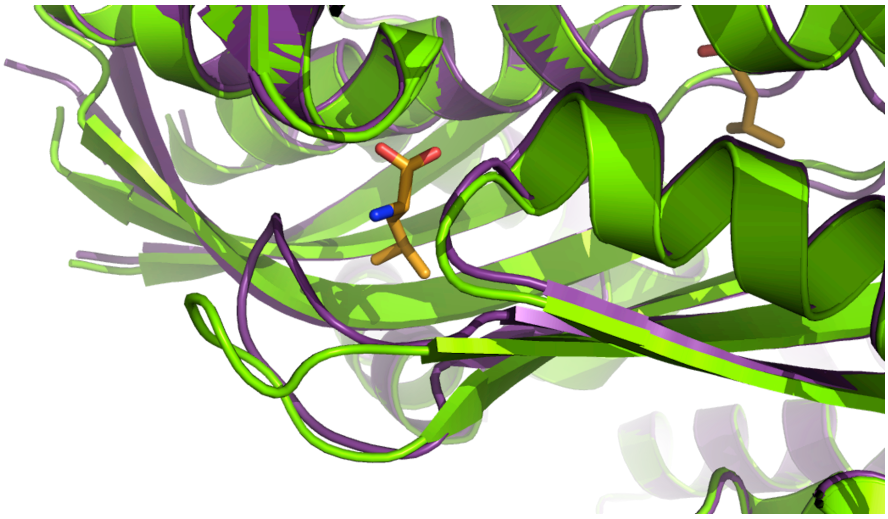


Figure 1.11: Loop closing in the leucine binding site of *Mtu*IPMS. Uninhibited structure 1SR9 is shown in green, with leucine-bound structure 3FIG in purple. Leucine is shown in orange.

Outside of the flexible loop, resistance to leucine may be bestowed by substitution of the conserved residue Gly479 for cysteine in *E. coli* α-IPMS (Gly622 in *Mtu*IPMS).[70] Resistance to leucine is also observed in the *S. cerevisiae* substitution of Ser519 for threonine (equivalent to leucine-binding residue Ala536 in *Mtu*IPMS), and enzyme activity is increased slightly in the presence of leucine in this variant.[69] Additionally, substitution of residues Glu540 or His541 in *S. cerevisiae* α-IPMS (Glu556 and His557 in *Mtu*IPMS) for lysine or proline, respectively, provides protection against CoA-mediated inactivation.

The leucine binding arrangement in *Mtu*IPMS is very similar to that observed for isoleucine in *L. interrogans* CMS (Figure 1.12). Here, isoleucine forms two hydrogen bonds from its carboxyl group to the main-chain amides of Asp431 and Gln495′ (Ala536 and Ile627′ in *Mtu*IPMS), whereas the isoleucine amino nitrogen shares three hydrogen bonds with the main-chain carbonyls of Thr464′, Pro493′ and Ala466′ (Asp563′, Pro625′ and Ala565′ in *Mtu*IPMS). The hydrophobic end of the inhibitor forms extensive van der Waals contacts with Tyr430, Leu451, Tyr454, Ile458′ and Val468′ (Leu535, Val551, Tyr554, Ala558′ and A567′ in *Mtu*IPMS). It appears that this hydrophobic pocket dictates enzyme specificity for isoleucine over leucine, as substitution of Val468 for alanine results in an enzyme that is inhibited by both of these branched-chain amino acids. There is little conservation between the α-IPMS and CMS inhibitor-binding sites, with only a tyrosine and a proline identical between them. The one residue in *Mtu*IPMS (Asn532) that contributes a hydrogen bond to leucine through its side-chain, rather than main-chain, functional groups is an aspartate in CMS, and is not in a position to interact with isoleucine.

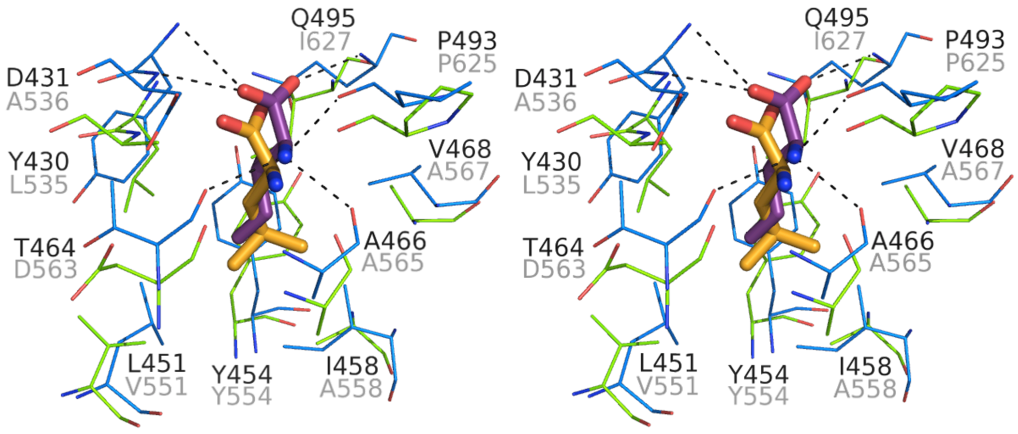


Figure 1.12: Stereo diagram of the isoleucine binding site in *L. interrogans* CMS (blue), overlaid with the leucine binding site of *Mtu*IPMS (green). Isoleucine is shown in purple, leucine in orange. Binding of isoleucine is indicated. Key residues are labelled for CMS (black) and α-IPMS (grey).

1.3 Protein allostery

Enzymes may be regulated at many points in their life cycle, from gene expression to protein degradation. Between these two extremes they may also be regulated by direct control of enzyme catalytic activity. Enzyme control can take many forms, but of particular interest is allosteric regulation, wherein action on one site of the protein may affect catalysis at a distant active site.

Two general theories of allostery were put forth in the 1960s: the Monod-Wyman-Changeux (MWC) model (also known as the concerted model) and the Koshland-Némethy-Filmer (KNF) model (also known as the sequential model). The MWC model suggests that all proteins exist in equilibrium between interconvertible tense and relaxed states, where the relaxed state is more receptive to substrate binding.[71] Binding of an allosteric inhibitor results in a shift in this equilibrium towards the tense state. The KNF model also posits that proteins have two conformational states, but that binding of an allosteric effector induces a conformational change. This induced change of state in one subunit then encourages allosteric ligand binding in other subunits in the oligmeric structure.[72] The subunits undergo conformational changes separately, in contrast to the MWC model, in which all subunits of an oligomer adopt the same state.

In addition to and expanding on these theories is the recent notion of "population shift", which states that proteins may exist in an ensemble of conformers due to their internal motions and flexibility.[73,74] Allosteric interactions may then cause a shift in this population to a new ensemble of states.

1.3.1 General mechanisms of allostery

Triggers of allosteric regulation include phosphorylation, formation or reduction of disulfide bonds, and binding of proteins or small molecules.[75] Binding of small molecules is the most common and is the form of regulation observed in α-IPMS.

The structural effects of regulator binding vary from large changes in protein structure, to movement of a single residue, to changes so subtle they are not apparent in structural analyses. At the more dramatic end of this spectrum is change in oligomeric state, such as that seen for *A. thaliana* TD, in which isoleucine causes a dissociation of the tetrameric enzyme to a dimeric form.[76] This inhibition and dissociation can be reversed by valine.[7] A change in oligomerisation is also seen when ATP phosphoribosyltransferase (ATP-PRTase) moves from an active dimer to an inactive

hexamer upon histidine binding. Here, the change in quaternary structure acts to block substrate access to the active site.[77] A smaller-scale gating of the active site is seen in 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase (DAH7PS) from *Thermotoga maritima*, in which tyrosine binding causes a movement of the regulatory domain to obstruct the active site, without otherwise affecting the tetrameric nature of the protein.[78]

Ligand binding may also affect the conformation of the active site itself. An isozyme of DAH7PS from *E. coli* conveys phenylalanine inhibition via the movement of four short peptides, culminating in changes to the active site that impair competent substrate binding.[79] In chorismate mutase the only noticeable change to the active site upon allosteric tyrosine binding is the insertion of a glutamate residue, which may be enough to alter the electrostatic properties of the substrate binding pocket.[80]

For enzymes without any obvious regulatory structural change, it is becoming increasingly apparent that regulation can be entirely driven by changes in dynamics. This has been clearly mapped in the dimeric catabolite activator protein, in which binding of cyclic AMP (cAMP) to one monomer enhances protein motions, negatively impacting binding of cAMP to the other monomer.[81]

Of course, these mechanisms are not mutually exclusive, and one protein may be affected by both local structural changes and collective motions in the conformational ensemble.

1.3.2 Molecular dynamics of *Mtu*IPMS inhibition

As mentioned above, there is very little difference between the leucine-bound and unbound *Mtu*IPMS X-ray crystal structures, so the mechanism of inhibition in this enzyme is unclear. Thus it is appealing to hypothesise that leucine inhibition is largely mediated by enzyme molecular dynamics. Solution-phase amide hydrogen/deuterium exchange (HDX) has been performed to elucidate possible changes in solvent accessibility between inhibited and uninhibited enzyme. This method measures the incorporation of deuterium into backbone amide hydrogens of proteolysed peptide fragments.[82] The incorporation can then be compared between different states of the protein to visualise changes that may be dynamic or artifactually suppressed in crystal structures.

A comparison between inhibited and uninhibited α-IPMS enzymes is shown in Figure 1.13. Deuterium incorporation is decreased in much of the regulatory domain and in

the region of subdomain II that interacts with this domain.[49] Decrease in exchange represents a shift in conformational equilibrium towards enhanced protection from solvent upon leucine binding.

Only one peptide in the catalytic domain showed a change in deuterium exchange upon leucine binding (residues 78–87), a peptide that includes α-KIV-binding residue Arg80 and metal-coordinating residue Asp81, as well as conserved residues Leu79, Gly82 and Gln84. This peptide also shows decreased deuterium exchange when conserved subdomain I residue Tyr410 is substituted with phenylalanine.[49] The Y410F variant retains the ability to bind leucine, with a K_D of $21 \pm 1\,\mu M$, yet has a catalytic rate similar to that of fully-inhibited wild-type *Mtu*IPMS, so it is thought to represent a constitutively-inhibited variant.[48]

While this experiment gives evidence on how feedback regulation may be transferred from the regulatory domain to subdomain II, and how it may affect key residues in the active site, the pathway between subdomain II and the site of catalysis remains unclear. This is due to a lack of change in subdomain I and peptides missing from the HDX analysis at the interface of the two subdomains.

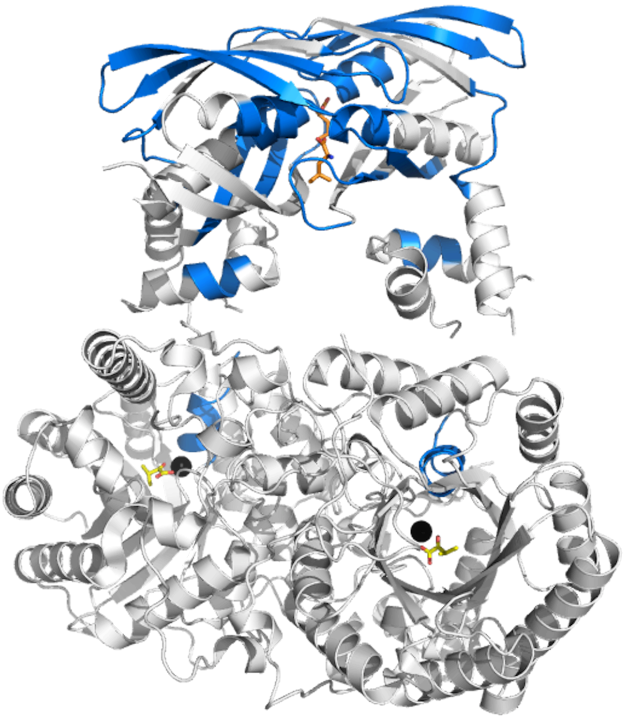


Figure 1.13: Map of changes in hydrogen/deuterium exchange in *Mtu*IPMS upon leucine binding. Areas of decreased deuterium incorporation are shown in blue. α-KIV is shown in yellow, Zn^{2+} in black and leucine in orange.

1.4 Objectives of this thesis

The broad objective of this thesis is to understand the nature of allosteric regulation in α-IPMS, including the effects of the regulatory domain itself on catalytic activity and the importance of structural asymmetry in the enzyme. This was investigated using α-IPMSs from the pathogens *M. tuberculosis* and *Neisseria meningitidis*.

More detailed goals of this research are:

- To expand the breadth of current α-IPMS knowledge by characterising a previously unstudied orthologue of this enzyme;
- To elucidate the importance of residues in the regulatory domain and subdomain II via the generation and characterisation of amino acid-substituted variants of the enzyme;
- To describe the contributions of the regulatory domain as a whole towards α-IPMS structure and function by characterisation of truncated variants;
- To probe the interfaces of the regulatory and catalytic halves of α-IPMS and evaluate their importance in protein asymmetry; and
- To determine whether the asymmetry seen in the *Mtu*IPMS crystal structures is also seen in solution.

α-IPMS from *N. meningitidis* (*Nme*IPMS) was selected for cloning and characterisation due to the pathogenicity of this organism. Amino acid-substituted variants of this enzyme were generated to investigate leucine-mediated regulation.

A truncated variant of *Nme*IPMS, lacking the regulatory domain, was cloned and compared to a similarly truncated variant of *Mtu*IPMS. Structural data found for truncated *Nme*IPMS were compared to existing *Mtu*IPMS data.

Lastly, residues at the domain interfaces of *Mtu*IPMS were substituted in an attempt to affect protein asymmetry. The solution-phase structures of these variants were compared to wild-type enzyme and the existing crystal structures.

Chapter 2

Wild-Type α-IPMS from *Neisseria meningitidis*

2.1 Overview

The first goal of this research was to clone and fully characterise wild-type α-IPMS from *N. meningitidis* (*Nme*IPMS). This characterisation included evaluation of protein size and oligomeric structure, and examination of general structural features of the enzyme in the presence of ligands (such as substrates or inhibitors).

The kinetics of this enzyme were elucidated, including the effect of pH on catalysis and the Michaelis-Menten constants for the catalytic substrates (AcCoA and α-KIV). The effects of the allosteric inhibitor leucine were also investigated.

Finally, the solution-phase structure of *Nme*IPMS was determined by small angle X-ray scattering (SAXS).

2.2 Background

N. meningitidis serotype B is the pathogen responsible for bacterial meningitis, a disease that can lead to death or permanent disfigurement.[83,84] The pathogenicity of this organism makes it an attractive target for antibiotic development. Research into the enzymes of *N. meningitidis* can thus aid future research into targeted inhibitors.

α-IPMS from *N. meningitidis* serotype B has a sequence similarity of 78 % compared with the characterised enzyme *Mtu*IPMS. *Nme*IPMS is the smaller of the two, with 517 residues and a monomer mass of 56 kDa, as opposed to the 644 residues and mass of 70 kDa for *Mtu*IPMS. These differences in size are partially accounted for by the presence of a 65 residue N-terminal extension in *Mtu*IPMS, that winds over the $(\beta/\alpha)_8$-barrel of the other monomer in the dimer (Figure 2.1). *Nme*IPMS also lacks the 38-residue tandem repeat found in the *Mtu*IPMS regulatory domain (residues 575–612; mostly undefined in structure).

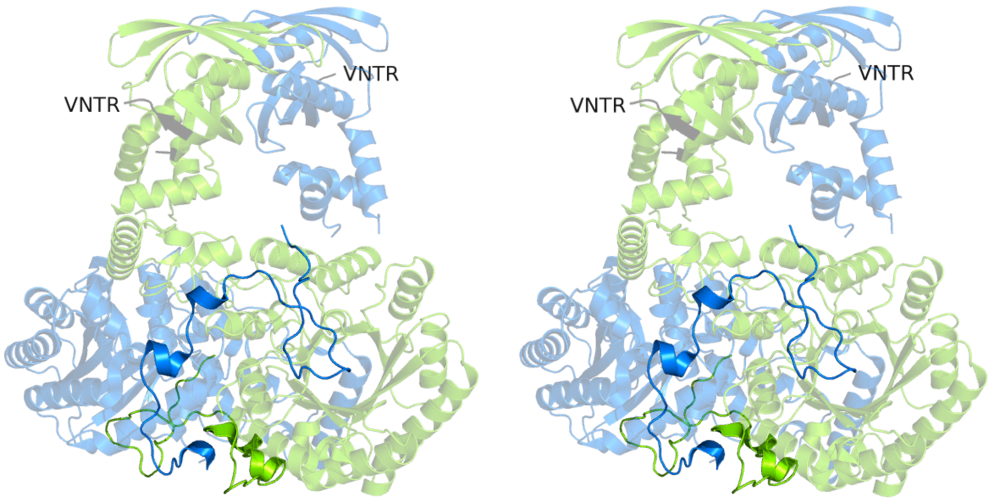


Figure 2.1: Stereo view of *Mtu*IPMS structure 1SR9 with N-terminal extension highlighted and position of the VNTR (mostly disordered) indicated in grey. Each monomer is shown in a different colour.

All key active-site residues in *Mtu*IPMS are conserved in *Nme*IPMS, including the metal-binding residues Asp81, His285 and His287 (Asp16, His204 and His206 in *Nme*IPMS), α-KIV-binding residues Arg80 and Thr254 (Arg15 and Thr173 in *Nme*IPMS), the potential catalytic base Glu218 (Glu143) and the two crossover residues from subdomain I, His379′ and Tyr410′ (His302′ and Tyr313′).

The leucine binding site is less well conserved, perhaps due to the fact that most hydrogen bonding to leucine observed in *Mtu*IPMS is through main-chain, as opposed to side-chain, functional groups. The leucine binding site is explored further in Chapter 3.

A homology model structure has been generated for *Nme*IPMS by Dr Wanting Jiao, based on the $(\beta/\alpha)_8$-barrel portion of a solved, truncated *Nme*IPMS structure (Section 4.2.7) and the subdomains and regulatory domain of *Mtu*IPMS structure 1SR9.[51] This homology structure compares to the *Mtu*IPMS crystal structure with a root mean squared (rms) difference of 1.7 Å for 659 Cα atoms (Figure 2.2). The modelled *Nme*IPMS structure lacks the *Mtu*IPMS N-terminal extension, and represents the complete *Nme*IPMS protein sequence except for a truncation of the final thirteen residues. These C-terminal residues did not align with any corresponding residues in *Mtu*IPMS, and were unable to be satisfactorily folded within the model.

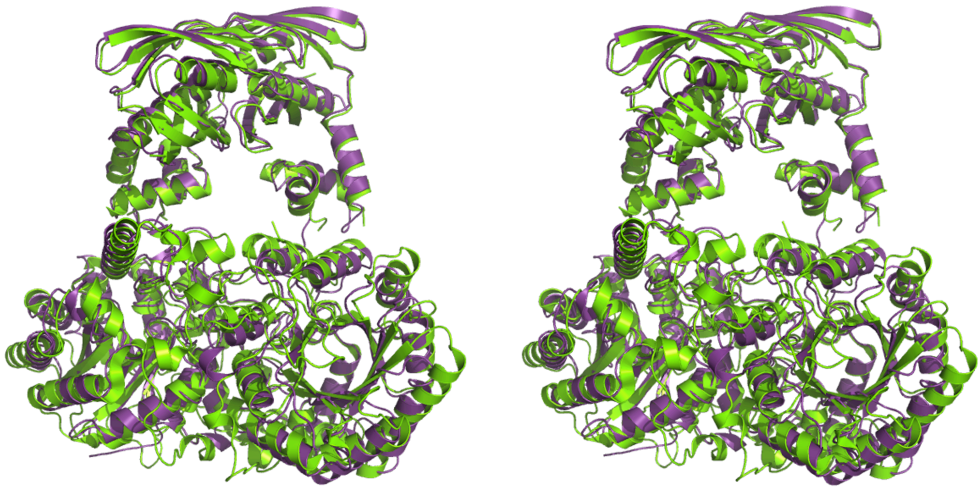


Figure 2.2: Homology model of *Nme*IPMS (purple) superimposed on the crystal structure of *Mtu*IPMS (PDB code 1SR9 — green).

2.3 Isolation of *Nme*IPMS

2.3.1 Cloning

Touchdown polymerase chain reaction (PCR) was performed to amplify the gene *leuA*, encoding α-IPMS, from *N. meningitidis* genomic DNA. Primers designed for PCR included a 5′ CACC overhang for topoisomerase-mediated ligation as described below.

The amplified PCR products were analysed by agarose gel electrophoresis, in which the *leuA* sample revealed a band appropriate for a calculated size of 1558 bases, including the 5′ overhang (Figure 2.3).

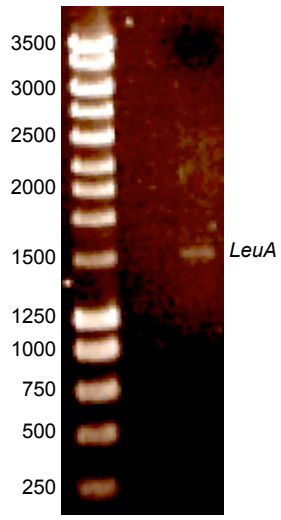


Figure 2.3: Agarose gel of cloned *leuA* DNA fragment. Marker sizes indicated in number of bases.

Topoisomerase was used to ligate the *leuA* gene into a pET-151 vector. The engineered 5′ overhang on the PCR fragment was ligated to the 3′ end of a region encoding a polyhistidine tag and a tobacco etch virus (TEV) protease cleavage sequence (detailed below in Figure 2.6).

DNA encoding the His-tag was appended to the *leuA* gene in the vector to allow the tagged protein to be easily purified. The TEV cleavage sequence allows recognition by TEV protease, allowing the targeted cleavage of the His-tag. A T7 promoter region binding a lac operon was also present in the plasmid, to enable over-expression of the gene of interest under high lactose conditions, and an ampicillin resistance

gene was included to easily select for cells possessing the plasmid by growth on ampicillin-containing media.

E. coli BL21 One Shot TOP10 cells were transformed with the *leuA*-containing construct, and plasmids were isolated from these cells for sequencing. Plasmid sequences were compared to gene sequences to confirm successful PCR and ligation, and plasmids that were shown to include the whole, correct *leuA* gene were transformed into *E. coli* BL21(DE3)Star cells for later protein expression.

2.3.2 Expression

Expression of *Nme*IPMS was carried out as described in Section 7.3.6. Cell pellets harvested after protein expression were either lysed immediately or stored at -80 °C until required. Steps from lysis to initial immobilised metal affinity chromatography (IMAC) were performed without delay to reduce the effects of proteolysis in the lysate. If samples of crude lysate were required to be stored for longer than an hour then protease inhibitors were added to the lysate.

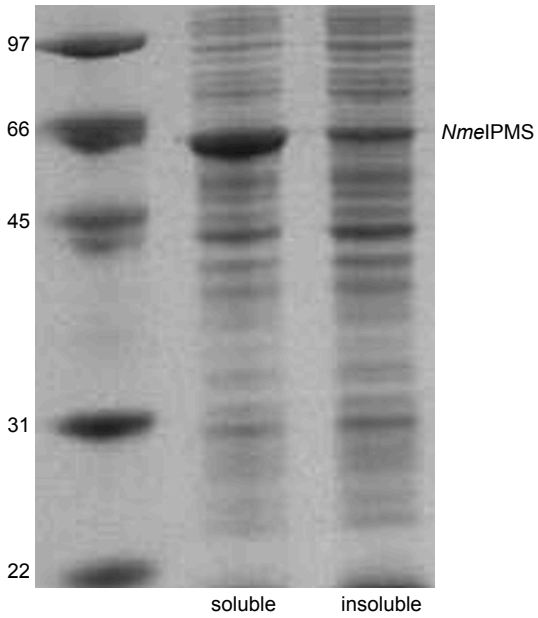


Figure 2.4: SDS-PAGE of *Nme*IPMS over-expression. Lanes show the soluble and insoluble fractions after cell lysis by sonication. Marker weights indicated in kDa.

Lysis was generally carried out by sonication, however larger cell cultures were occasionally lysed by a high-pressure cell disruptor. Figure 2.7 shows a sodium-dodecyl-

sulfate polyacrylamide gel electrophoresis (SDS-PAGE) analysis of proteins in the soluble and insoluble fractions after protein expression and cell lysis by sonication. The majority of over-expressed protein is found in the soluble fraction, at the expected weight of ~60 kDa (including the 4 kDa His-tag).

2.3.3 Purification

Typical purification followed the steps outlined in Figure 2.5.

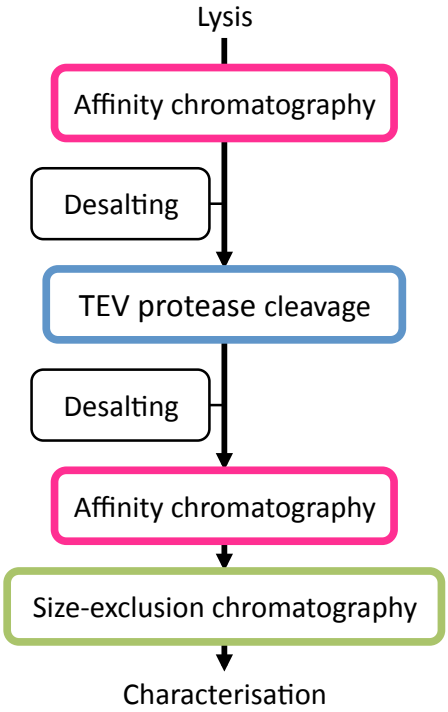


Figure 2.5: Flow diagram of *Nme*IPMS purification procedure

This protocol resulted in protein of sufficient purity for all characterisation assays. The first desalting step removes high concentrations of imidazole from the protein solution in preparation for overnight TEV protease cleavage, while the second desalting step removes dithiothreitol (DTT) and ethylenediaminetetraacetic acid (EDTA) from the cleavage buffer to allow for further affinity chromatography.

This procedure typically gave a final protein yield of 40 mg for 1 L of lysogeny broth (LB) cell culture.

Immobilised metal affinity chromatography

IMAC was the first step in protein purification. Co^{2+}-bound Talon® affinity resin was used to isolate *Nme*IPMS from background *E. coli* proteins via binding of the engineered N-terminal His-tag. Bound protein was eluted with 150 mM imidazole. Typical unbound and bound fractions from this process can be seen in lanes 2 and 3 of Figure 2.7.

To improve the yield of tagged protein from this step, fractions of protein that did not bind to the column were often reloaded onto the resin (following re-equilibration in buffer with a low imidazole concentration) and a second elution performed. This step allowed for isolation of tagged protein that may not have bound to the resin during the first run.

Desalting

*Nme*IPMS was eluted from the Talon® column in buffer with a high imidazole concentration. Pooled elution fractions were thus desalted into a low-imidazole buffer at this stage to reduce any effects of imidazole on stability and TEV protease efficacy.

TEV protease digestion

Once in buffer with a reduced imidazole concentration, protein was digested at 4 °C overnight with TEV protease to cleave the His-tag from *Nme*IPMS, minimising the potential effects of this tag on protein structure and function.

The His-tag added 33 amino acid residues to the N-terminus of *Nme*IPMS, including six histidine residues, to bind affinity resin, and a TEV protease recognition sequence (Figure 2.6). After digestion, the six residues GIDPFT from this tag remained on the N-terminus of *Nme*IPMS.

MHHHHHHGKPIPNPLLGLDSTENLYFQGIDPFT

Figure 2.6: Sequence of vector pET-151 polyhistidine tag, detailing: histidine binding region (pink); TEV protease cleavage site (blue); and post-cleavage N-terminal protein extension (orange).

After digestion, the protein-containing solution was again desalted, as above, to remove the DTT and EDTA present in the digestion buffer. The protein was then re-applied to the Talon® affinity column, which this time bound the isolated tag and the TEV protease (itself possessing a His-tag) and separated them from the newly untagged *Nme*IPMS.

Lane 4 of the gel shown in Figure 2.7 shows *Nme*IPMS after TEV digestion and the second IMAC separation. The single band at a decreased size from the protein pre-TEV addition indicates that *Nme*IPMS has been cleaved. A faint band at the pre-TEV protein weight can be seen in the column-bound fraction (lane 5) and indicates that TEV cleavage was not 100 % efficient. TEV protease itself is also visible as a faint band at 28 kDa in the column-bound fraction.

Size-exclusion chromatography

The final step in purification of *Nme*IPMS was size-exclusion chromatography (SEC), from which *Nme*IPMS eluted as a single peak. The weight of the purified protein after this step is consistent with the calculated mass of 56 kDa (Figure 2.7).

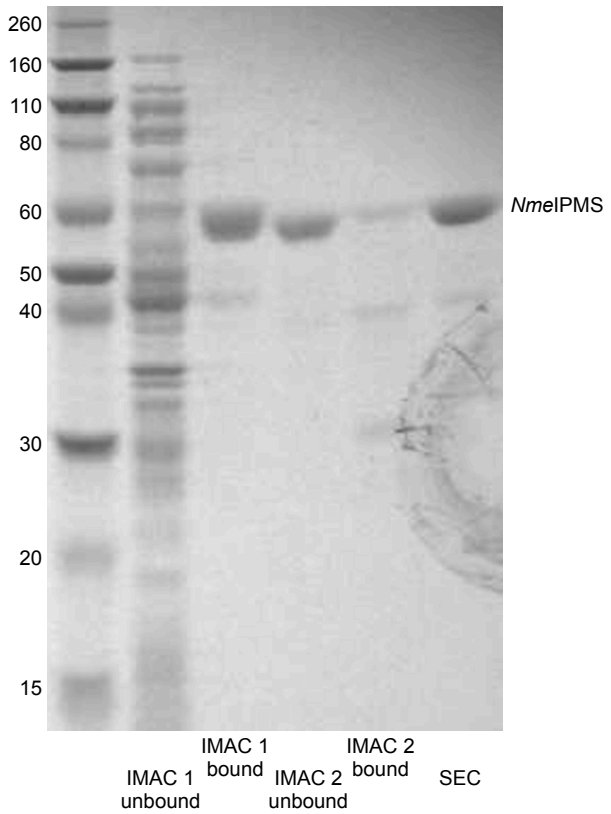


Figure 2.7: SDS-PAGE of *Nme*IPMS purification. Marker weights indicated in kDa.

2.4 Physical characterisation

2.4.1 Presence of secondary structure

A circular dichroism (CD) spectrum (Figure 2.8) was obtained to determine whether purified *Nme*IPMS was properly folded.

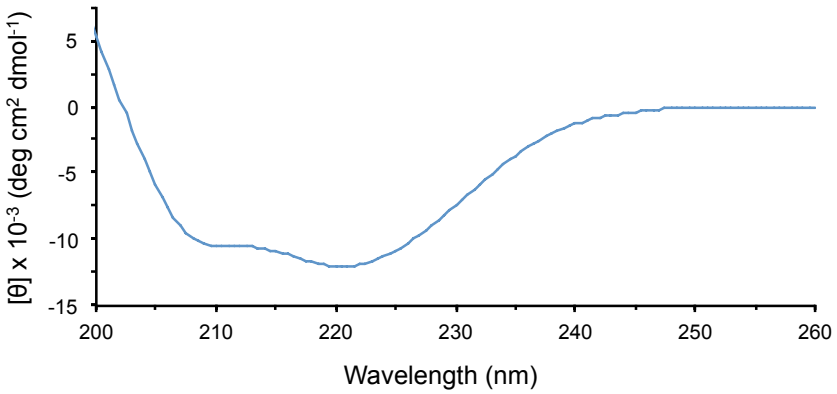


Figure 2.8: Circular dichroism spectrum of *Nme*IPMS

Analysis of these data on the K2D3 secondary structure server[85] estimates that the protein has a ratio of α-helix to β-strand of 1.09. This is about half the ratio expected from the α-IPMS homology model and could indicate some disorder in the α-helical regions of the protein, but more likely is a reflection of the highly variable nature of secondary structure estimation.[86] The presence of secondary structure in the protein indicates that it is folded and suitable for further characterisation.

2.4.2 Molecular mass

The molecular weight of *Nme*IPMS was measured by electrospray mass spectrometry (MS). The mass for this protein was found to be 56029 Da, which compares favourably with the mass calculated from the amino acid sequence by ProtParam[87] (56027 Da) and the size of the band observed in SDS-PAGE (Figure 2.7).

2.4.3 Oligomeric structure

Many α-IPMS enzymes are reported to be dimers in solution, although some may form trimers or tetramers.[41,42,46,65] The multimeric state of *Nme*IPMS in solution was determined using an analytical gel-filtration column.

A standard curve was generated from proteins of known size (Table 2.1). Blue dextran was large enough to elute at the void volume (V_o) of the column. The standard curve was calculated from a plot of the log of protein molecular weight vs. the ratio of elution volume (V_e) to void volume. A line of best fit for this curve was calculated as $y = -1.6x + 7.8$ with $R^2 = 0.98$ (Figure 2.9).

Protein	Molecular Weight (kDa)	V_e (mL)
Carbonic Anhydrase	29	16.6
Ovalbumin	43	15.3
Bovine Serum Albumin	66	14.1
Conalbumin	75	14.4
Alcohol Dehydrogenase	150	13.1
Ferritin	440	10.4
Apoferritin	443	11.0
Thyroglobulin	669	9.5
Blue Dextran	2000	8.2

Table 2.1: Protein standards used in analytical gel-filtration chromatography

*Nme*IPMS eluted as a single peak at 13.3 mL. The equation of the standard curve was used to convert this to a molecular weight of 115 kDa (Table 2.2). The calculated molecular weight of monomeric *Nme*IPMS is 56 kDa, so this protein appears to be a dimer in solution.

Effects of leucine on structure

Crystal structures of *Mtu*IPMS with leucine bound have very little structural difference from the non-leucine bound enzyme structure.[51] The flexible loop over the leucine binding site becomes more ordered, but there is no observed change to the catalytic site and no change in oligomeric structure.

To eliminate the possibility that *Nme*IPMS undergoes changes in quaternary structure upon leucine binding, further analytical gel-filtration chromatography was car-

ried out in the presence 200 µM L-leucine (Figure 2.9 and Table 2.2). *Nme*IPMS again eluted at a volume consistent with the mass of the dimer (117 kDa). This enzyme has residual activity of less than 50 % at this leucine concentration (see Section 2.5.2 below).

*Nme*IPMS, like *Mtu*IPMS,[46] appears to be a dimer in solution in the presence and absence of the allosteric inhibitor. This precludes the possibility that inhibition is effected by gross changes in oligomeric structure, such as the removal of subdomain I residues from the active site by dissociation of the dimer. It is possible, however, that at higher or lower protein concentrations different behaviour could be observed.

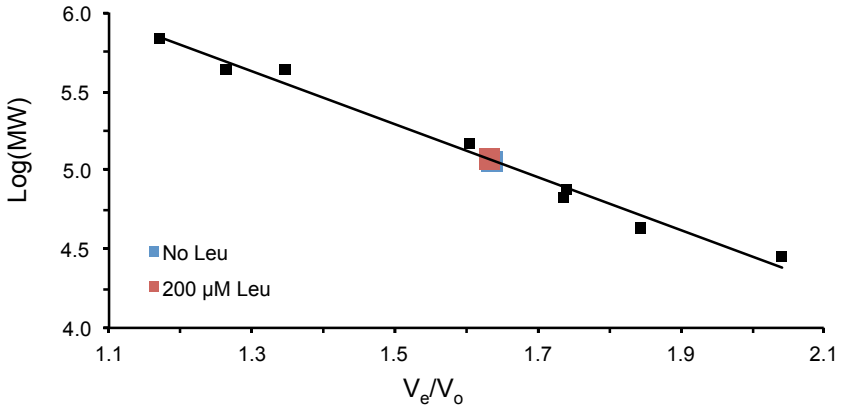


Figure 2.9: *Nme*IPMS gel filtration in the presence and absence of 200 µM L-leucine. Samples run in 50 mM TrisHCl buffer (pH 7.5) with 100 mM KCl. The calibration curve is shown in black.

	Molecular Weight (kDa)		
	Monomer[a]	Oligomer[b]	Quaternary Structure
*Nme*IPMS	56	115	dimer
*Nme*IPMS + leu	56	117	dimer

Table 2.2: Oligomeric structure of *Nme*IPMS in the presence and absence of 200 µM L-leucine. Samples run in 50 mM TrisHCl buffer (pH 7.5) with 100 mM KCl. [a]Calculated based on sequence; [b]Determined from calibration curve.

2.4.4 Thermal stability

Differential scanning fluorimetry (DSF) was used to investigate the stability of *Nme*-IPMS in the presence of substrates and inhibitor. DSF uses a dye that fluoresces in the presence of hydrophobic amino acids to monitor protein denaturation. As the protein is heated, the internalised, hydrophobic residues are exposed to the dye, and fluorescence increases. Thus the temperature of denaturation can be determined as the point of greatest increase in fluorescence (Figure 2.10).

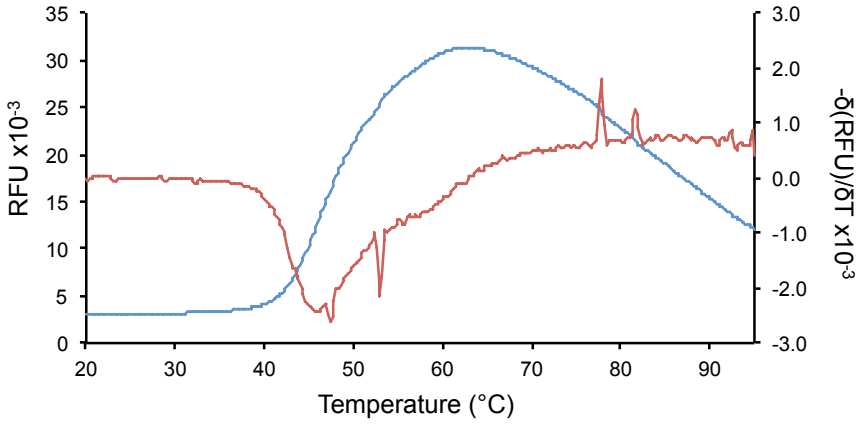


Figure 2.10: Example of a thermal denaturation curve produced by DSF, where *RFU* is *relative fluorescence units*. The negative of the first derivative is shown in red on the secondary axis. T_m is indicated by the point of greatest increase in fluorescence, or the minimum of the derivative.

DSF shows that *Nme*IPMS has a denaturation temperature of 44.5 ± 0.1 °C (Figure 2.11 and Table 2.3). This temperature is low compared to the *N. meningitidis* growth environment of 37 °C and the *in vitro* activity optimum of 40 °C for this enzyme.[63] It is likely that other factors present in the *N. meningitidis* cell increase stability, and other studies have shown that metal ions may contribute to this (see Section 2.7.1).[63] Noticeable increases in the denaturation temperature of *Nme*IPMS are observed in the presence of the substrate α-KIV or the inhibitor leucine (Figure 2.11). Leucine in combination with either substrate confers greater stability than either ligand alone. AcCoA has a minimal effect on this enzyme stability.

DSF can also be used as an indicator of ligand binding, if it is assumed that such binding has an effect on overall protein stability. The effects of amino acids isoleucine, valine and glycine were thus compared to that of leucine to give some measure of inhibitor specificity (Figure 2.12). While increasing concentrations of leucine give steadily increasing denaturation temperatures, isoleucine and glycine show no

significant effects while valine confers a modest increase in stability only at high concentration.

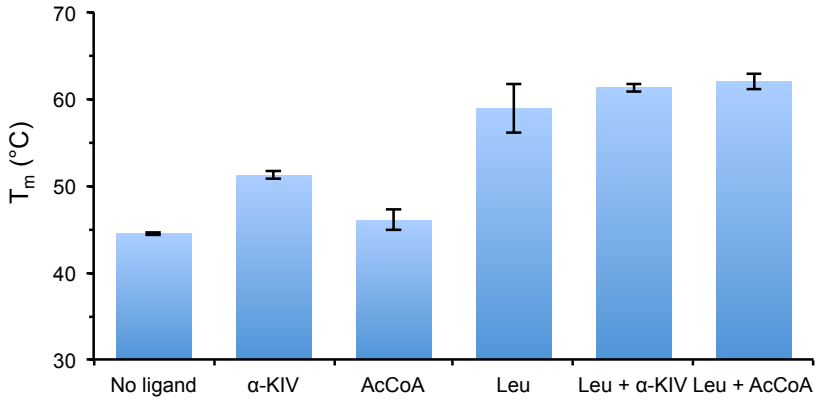


Figure 2.11: Denaturation temperatures of *Nme*IPMS in the presence and absence of biological substrates (250 µM each) and inhibitor (1 mM). All samples in 25 mM BTP buffer (pH 7.0).

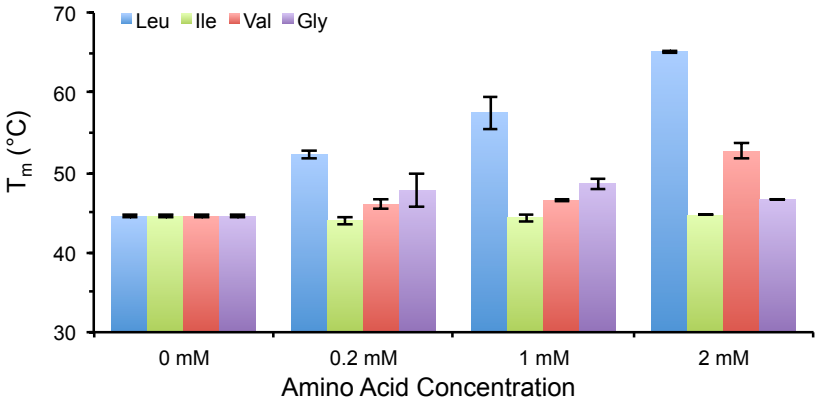


Figure 2.12: *Nme*IPMS denaturation temperatures in the presence of amino acids L-leucine, L-isoleucine, DL-valine and glycine. All samples in 25 mM BTP buffer (pH 7.0).

Ligand(s)	T_m (°C)
No ligand	44.5 ± 0.1
biological ligands	
250 µM α-KIV	51.2 ± 0.4
250 µM AcCoA	46.1 ± 1.1
1 mM L-leucine	59 ± 3
250 µM α-KIV + 1 mM L-leucine	61.3 ± 0.5
250 µM AcCoA + 1 mM L-leucine	62.0 ± 0.9
amino acids	
200 µM L-leucine	52.2 ± 0.5
1 mM L-leucine	59 ± 3
2 mM L-leucine	65.2 ± 0.2
200 µM L-isoleucine	43.9 ± 0.4
1 mM L-isoleucine	44.3 ± 0.5
2 mM L-isoleucine	44.6 ± 0.1
200 µM DL-valine	46.0 ± 0.5
1 mM DL-valine	46.5 ± 0.1
2 mM DL-valine	53 ± 1
200 µM glycine	48 ± 2
1 mM glycine	48.6 ± 0.7
2 mM glycine	46.6 ± 0.1

Table 2.3: Denaturation temperatures (T_m) of *Nme*IPMS in the presence of potential ligands. All samples in 25 mM BTP buffer (pH 7.0).

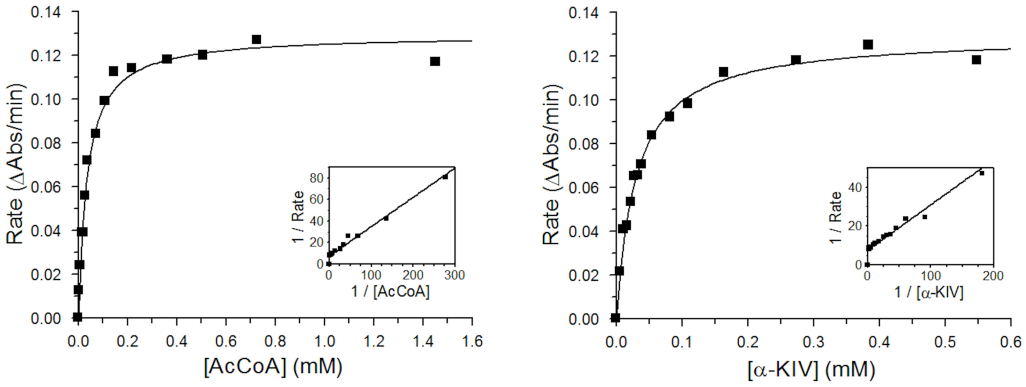
2.5 Kinetic characterisation

The activity of *Nme*IPMS was studied in two parts: first the catalytic constants were calculated for the two substrates (α-KIV and AcCoA), and then the effects of leucine on catalysis were investigated.

The majority of these kinetic assays were carried out using a 4,4′-dithiodipyridine (DTP)-coupled assay, monitoring absorbance at 324 nm. This measured the generation of a thio-pyridine by reaction of DTP with the free thiol of the enzyme product CoA.

2.5.1 Michaelis-Menten kinetics

The turnover number and apparent K_m values were found for *Nme*IPMS at 25 °C (Figure 2.13). Details of these data can be found in Table 2.4. These values show that *Nme*IPMS has a lower affinity for α-KIV and a higher affinity for AcCoA than *Mtu*IPMS.[46]



(a) AcCoA Michaelis-Menten plot **(b)** α-KIV Michaelis-Menten plot

Figure 2.13: Plots of kinetic data for *Nme*IPMS, showing variation in initial reaction rate with change in substrate concentration. *Inset:* Lineweaver-Burk plots of data.

*Nme*IPMS was also found to hydrolyse AcCoA independently of α-KIV. This uncoupled hydrolysis occurred at a rate 1000-fold slower than the condensation reaction, and had almost 10-fold decreased affinity for the substrate. These kinetic parameters are similar to those found for this reaction in *Mtu*IPMS.[46] The uncoupled hydrolysis of AcCoA in *Nme*IPMS was not affected by L-leucine at concentrations up to 10 mM.

	*Nme*IPMS	*Mtu*IPMS[46]
normal catalysis		
K_m^{app} for α-KIV (μM)	30 ± 2	12 ± 1
K_m^{app} for AcCoA (μM)	35 ± 3	136 ± 5
k_{cat} (s^{-1})	12.8 ± 0.3	3.5 ± 0.1
uncoupled AcCoA hydrolysis		
K_m^{app} for AcCoA (μM)	250 ± 30	160 ± 29
k_{cat} (s^{-1})	0.011 ± 0.001	0.03 ± 0.002

Table 2.4: Kinetic data for *Nme*IPMS and *Mtu*IPMS.[46] For *normal catalysis*, invariant substrate was held at a concentration of 250 μM. *Uncoupled AcCoA hydrolysis* indicates values obtained in the absence of α-KIV.

2.5.2 pH dependence

*Nme*IPMS activity varies with pH, with maximal activity observed at pH 8.5 (Figure 2.14). Assays were carried out using a direct assay at 232 nm, measuring the loss of AcCoA, as high pH caused the reaction of DTP and CoA to become rate limiting. Due to high AcCoA absorbance at 232 nm, non-saturating (80 μM) concentrations of this substrate were used. de Carvalho *et al.* have previously established that *Mtu*-IPMS has maximal k_{cat} values from pH 6.7–9.0 using the DTP-coupled assay.[46]

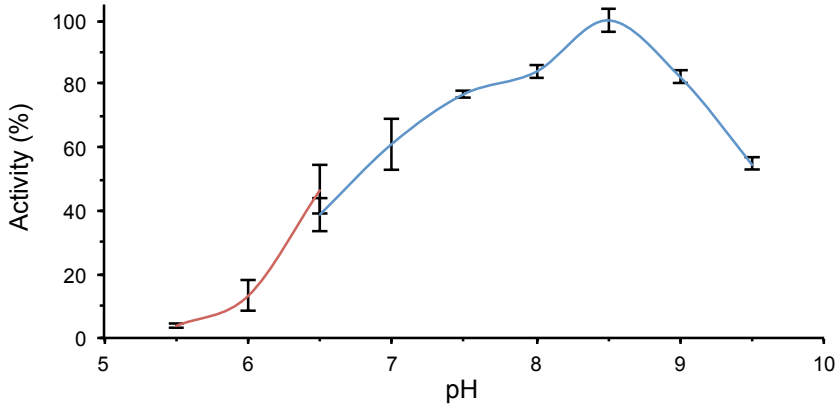
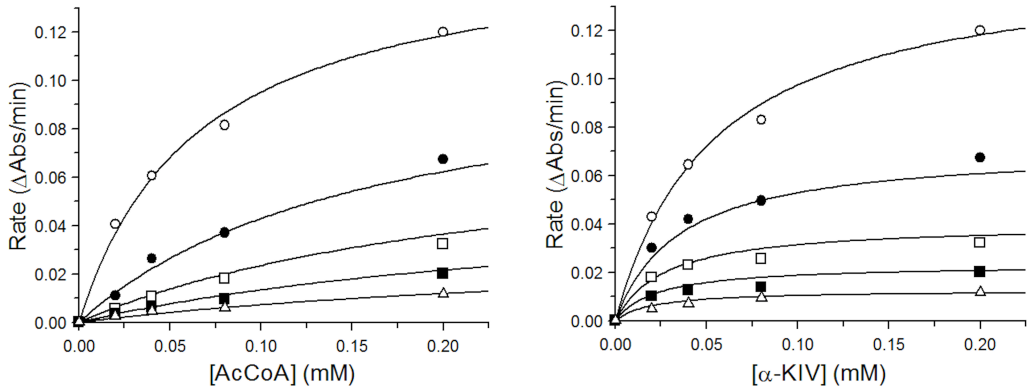


Figure 2.14: Effects of pH on *Nme*IPMS catalytic activity. Residual activity is normalised to the maximal acitivity seen (pH 8.5). Samples in BTP (blue) or MES buffer (red).

2.5.3 Allosteric inhibition

*Nme*IPMS, like *Mtu*IPMS, exhibits inhibition by leucine, the end-product of the α-IPMS metabolic pathway. Leucine inhibition of *Nme*IPMS was characterised with respect to inhibition constants, time dependence and pH dependence.

Kinetic parameters



(a) AcCoA inhibition plot **(b)** α-KIV inhibition plot

Figure 2.15: Plots for leucine inhibition of *Nme*IPMS. L-Leucine concentrations (from top to bottom) are 0 µM, 20 µM, 50 µM, 100 µM and 200 µM. 250 µM of each substrate was used.

In *Nme*IPMS, inhibition by leucine fits a mixed, non-competitive model with respect to both α-KIV and AcCoA (Figure 2.15), with K_i and K_i' values as indicated in Table 2.5. These data show that leucine inhibition is more acute with respect to AcCoA than α-KIV, and that the affinities of leucine for free and substrate-bound enzyme are very similar. These affinities are also very similar to those found for *Mtu*IPMS.[45]

	*Mtu*IPMS[45]	*Nme*IPMS	
	α-KIV	α-KIV	AcCoA
K_i (µM)	22 ± 2	57 ± 23	8 ± 1
K_i' (µM)	8 ± 1	18 ± 2	52 ± 17

Table 2.5: L-Leucine inhibition data for *Nme*IPMS

Some inhibition of *Nme*IPMS catalytic activity was also observed for high concentrations of valine, with 1 mM DL-valine causing a drop in activity of 6 %. No inhibition was observed for L-isoleucine or glycine at concentrations of up to 2 mM. It appears that feedback inhibition in this enzyme is highly selective in its amino acid sensitivity.

Time dependence

One of the noteworthy features of the leucine-mediated inhibition of *Mtu*IPMS is the time-dependency observed. In the presence of L-leucine, reactions catalysed by *Mtu*IPMS showed an initial burst of product formation, followed by a slower, linear rate.[45] This time dependence was not evident for leucine inhibition of *Nme*IPMS. CoA formation by *Nme*IPMS was measured, via its reaction with DTP, in the presence of saturating concentrations of substrates and metal ions, and 0–500 µM L-leucine. No obvious change in CoA production rate can be seen over a 9 min reaction period (Figure 2.16), which is in contrast to results reported for *Mtu*IPMS, where a noticeable decrease in enzyme activity over the course of a few minutes was observed.[45]

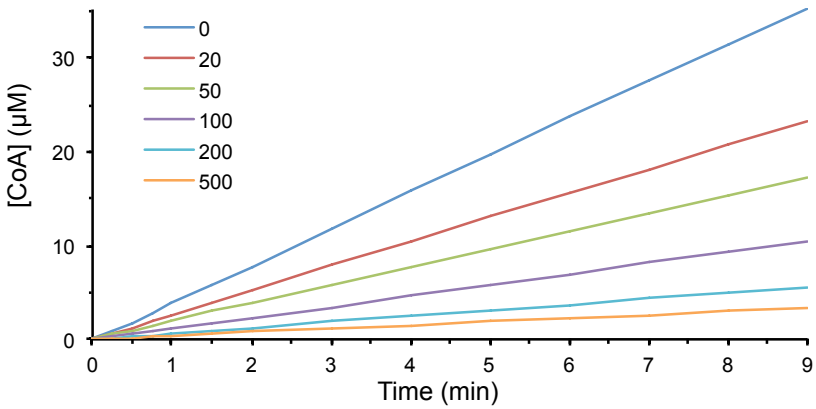


Figure 2.16: Leucine inhibition of *Nme*IPMS over time. Legend shows leucine concentration in µM. 250 µM of each substrate was used.

The transition in *Mtu*IPMS from a Enz·Leu complex to a more tightly bound Enz*·Leu complex is mediated by the flexible loop covering the leucine binding site.[44] Aside from a conserved glycine, the residues in this loop are poorly conserved between species (Figure 2.17), and it is possible that slow-onset regulation in α-IPMS is unique to *M. tuberculosis*.

Figure 2.17: Conservation of residues in the regulatory loop. Absolutely conserved residues are shown in red, strongly conserved residues in orange, and conserved branched-chain amino acids in green. Boxed residues are those found in the *Mtu*IPMS regulatory loop.[44,51]

pH dependence

The inhibition of *Nme*IPMS by leucine varies with pH, with the enzyme exhibiting enhanced inhibitor sensitivity in more acidic enrivonments (Figure 2.18). Assays were carried out at pH 6.5, 7.5 and 8.5, with concentrations of L-leucine from 0–2 mM. 200 µM inhibitor caused a loss of 20 % maximal *Nme*IPMS activity at pH 8.5, but a loss of 89 % activity at pH 6.5. At the highest concentration assayed (2 mM), leucine reduced activity by 50 % at pH 8.5 and 95 % at pH 6.5.

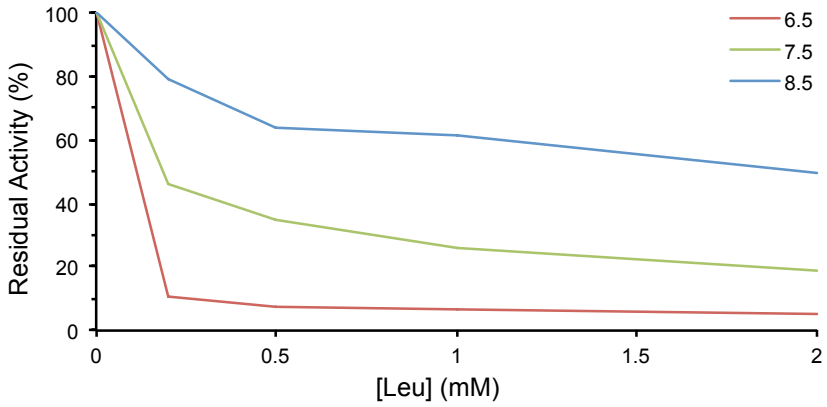


Figure 2.18: pH dependence of *Nme*IPMS L-leucine inhibition at 80 µM AcCoA and 250 µM α-KIV. Legend indicates pH.

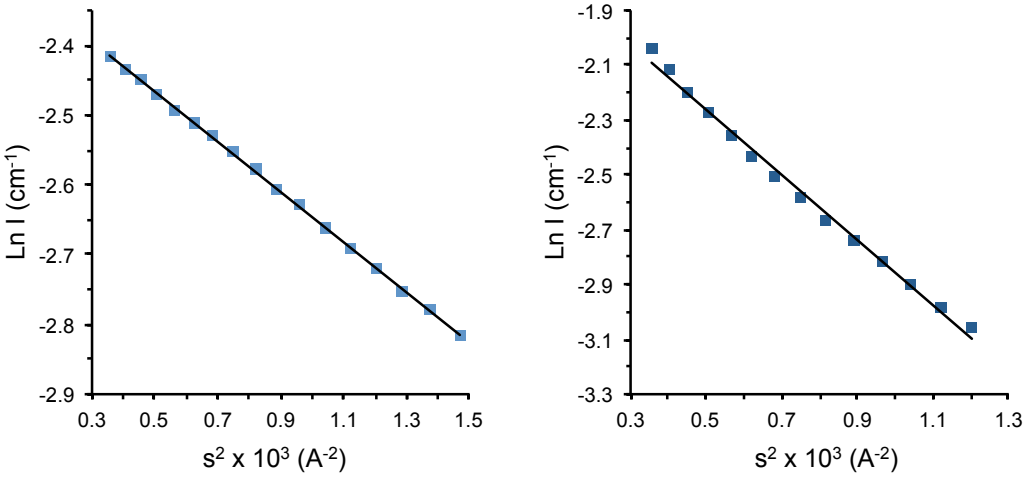
Such pH dependence of allosteric regulation has also been established for α-IPMSs from *Saccharomyces sp.*[58] and *S. typhimurium*,[53] the latter of which was found to be 30-fold more sensitive to leucine inhibition at pH 6.5 than at pH 8.5.

2.6 Small angle X-ray scattering

SAXS data were collected for *Nme*IPMS in the presence and absence of 1 mM L-leucine. These data were compared to investigate potential changes in protein solution-phase structure upon inhibitor binding.

2.6.1 Sample validation

Aggregation of samples was assessed by examining Guinier distributions of the data (Figure 2.19). These plots indicate aggregation by the presence of a non-linear dependence of the log of intensity vs. s^2. Unfortunately the leucine-bound sample showed signs of aggregation on this plot and no further interpretation of data was performed. Estimates of the radius of gyration (R_g) and the intensity at $s = 0$ ($I(0)$) were calculated from this distribution (Table 2.6).



(a) *Nme*IPMS **(b)** *Nme*IPMS with L-leucine

Figure 2.19: Guinier distributions of *Nme*IPMS SAXS data

2.6.2 Structure parameters

A pair-distribution function ($P(r)$ — Figure 2.20) was generated for the ligand-free *Nme*IPMS data. R_g and $I(0)$ were calculated from this plot and compared well with those calculated from the Guinier distribution. The agreement between

these two sets of values is another indication that the sample was not affected by aggregation.

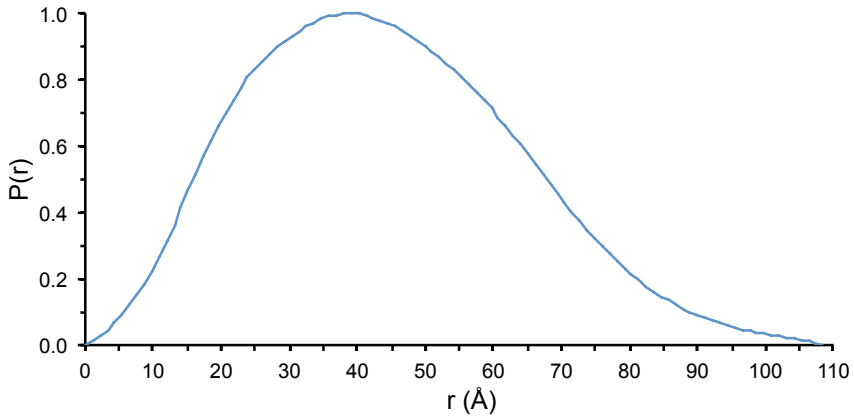


Figure 2.20: Pair-distribution function of *Nme*IPMS SAXS data. $P(r)$ function normalised to 1.

The $P(r)$ distribution also allowed for the calculation of the maximum dimension of the molecule ($D_{\max}$) and the Porod estimate of excluded volume (Table 2.6). This latter value is approximately proportional to the particle molecular mass, and this indicates that the protein is a dimer in the sample (Table 2.6). This is in agreement with the results from analytical gel-filtration experiments.

	*Nme*IPMS
structural parameters	
$I(0)$ (cm^{-1})	0.104 ± 0.004
from Guinier	0.102 ± 0.000
R_g (Å)	34.1 ± 0.1
from Guinier	32.9 ± 0.2
$D_{\max}$ (Å)	108.5
Porod volume estimate ($\mathrm{Å}^3$)	186000
molecular mass estimate	
mass estimate from Porod[a] (Da)	116000
monomeric mass from sequence (Da)	56000
likely oligomeric structure	dimer

Table 2.6: SAXS parameters for *Nme*IPMS. Data collected at a wavelength of 1.127 Å. [a]Particle mass is roughly related to Porod volume by a factor of 0.625.[88]

2.6.3 Comparison to homology model

Theoretical scattering patterns from the *Nme*IPMS homology model dimer and each of its constituent monomers were fitted to the experimental data with CRYSOL[89] (Figures 2.21 and 2.22). Fitting both dimer and monomer structures allows additional confirmation of the oligomeric structure present in the SAXS experiment. Theoretical scattering from each monomer was fitted separately, as each monomer adopts a markedly different conformation in the asymmetric crystal structure. The best fit to the experimental data was observed for the dimer, confirming that this is the oligomeric structure observed in solution. The similarity between the theoretical and experimental scattering patterns suggest that the *Nme*IPMS homology model dimer is a good representation of the solution-phase structure of the protein.

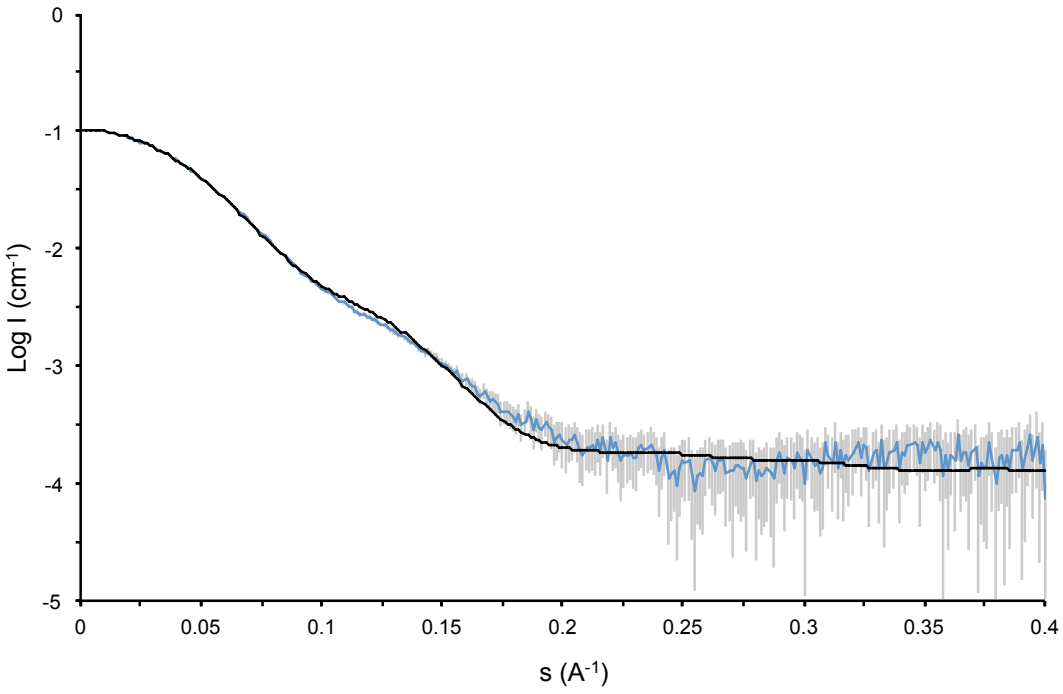


Figure 2.21: Fit of theoretical scattering from homology model dimer to *Nme*IPMS SAXS data. Fit from the model dimer is shown in black ($\chi = 1.16$). Experimental data are shown in blue, with error bars in grey.

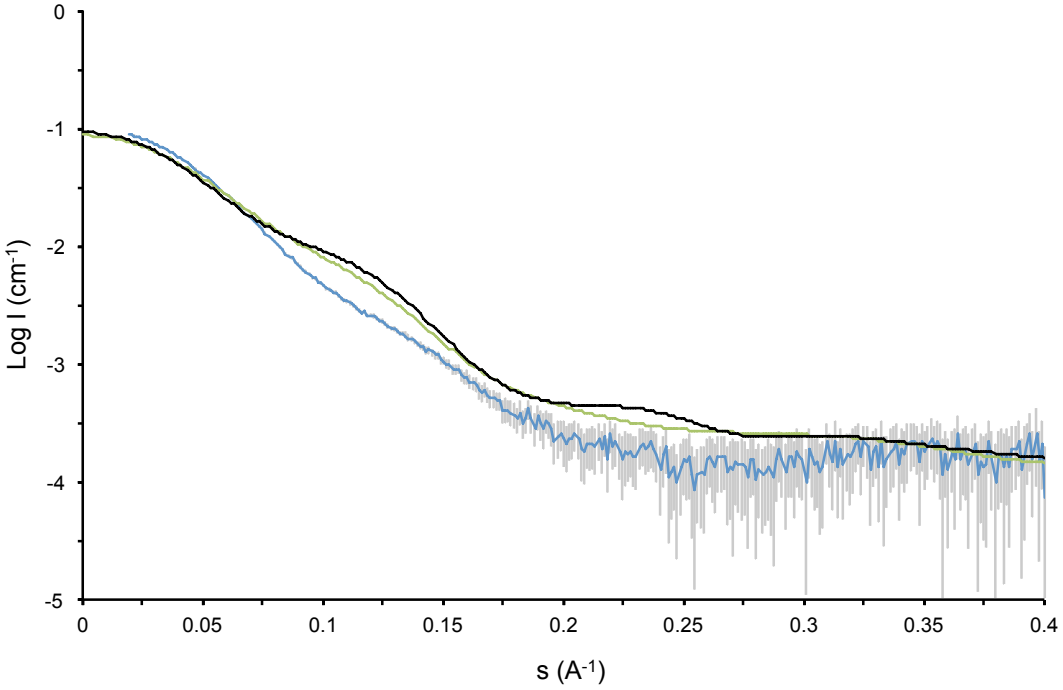


Figure 2.22: Fit of theoretical scattering from homology model monomers to *Nme*IPMS SAXS data. Fits of each monomer are shown in black ($\chi = 13.1$) and green ($\chi = 10.9$). Experimental data are shown in blue, with error bars in grey.

2.6.4 Theoretical domain movement in *Nme*IPMS

Due to the asymmetry in the α-IPMS dimer, it is tempting to postulate a hypothetical movement of the C-terminal domains relative to the N-terminal barrel. This could take the form of an oscillation from one asymmetric state to the other, in which the monomers have exchanged conformations (Figure 2.23). Such an oscillation would represent a movement of up to 80 Å for individual residues in *Mtu*IPMS, and would be energetically unfavourable unless it conferred some advantage on the enzyme. Movement between these conformational extremes or to a more symmetrical state could play a role in α-IPMS catalysis or feedback inhibition.

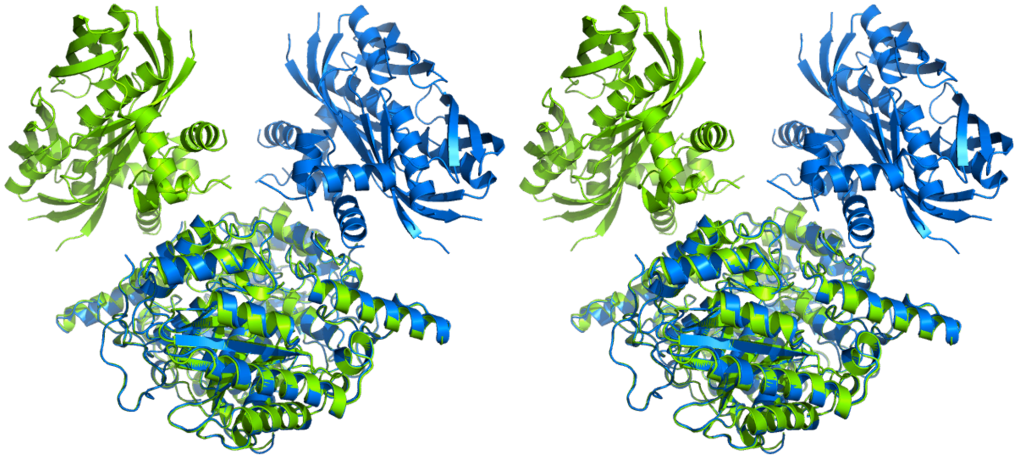
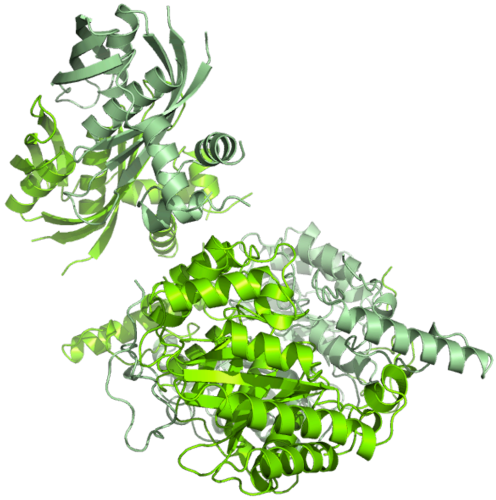
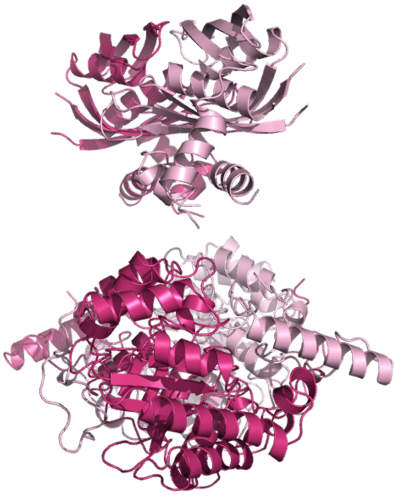


Figure 2.23: Stereo view of the potential movement of the α-IPMS regulatory domain. Overlay of two dimers of *Mtu*IPMS, rotated 180° relative to each other and aligned by $(\beta/\alpha)_8$-barrels.

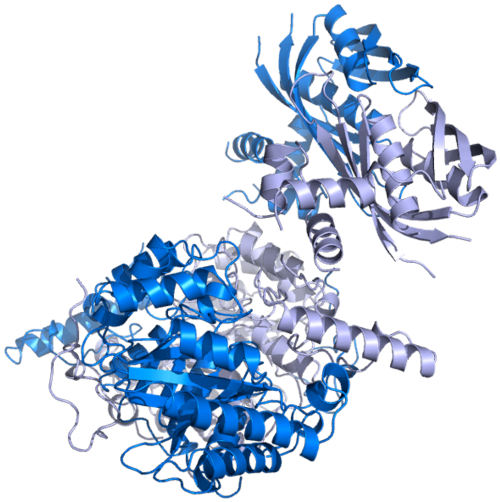
A simulated model of domain movement was created to compare with experimental SAXS data, and perhaps provide insight into the solution-phase structure of the protein. Twenty model structures were generated, representing the movement from the structure seen in the homology model to the diametrically opposite dimer conformation. The end-point structures and a symmetrical mid-point structure are shown in Figure 2.24. Theoretical SAXS curves were generated for each of these twenty models (Figure 2.25), but these show little difference between the structures. Therefore, this technique cannot elucidate the asymmetry present in *Nme*IPMS in solution.



(a) Model1 — Initial structure



(b) Model10 — Mid-point of potential movement



(c) Model20 — End-point of potential movement

Figure 2.24: Morph models of the potential movement of α-IPMS regulatory domain. Three model dimers from the twenty generated across the movement are shown. Each monomer is shown in a different colour.

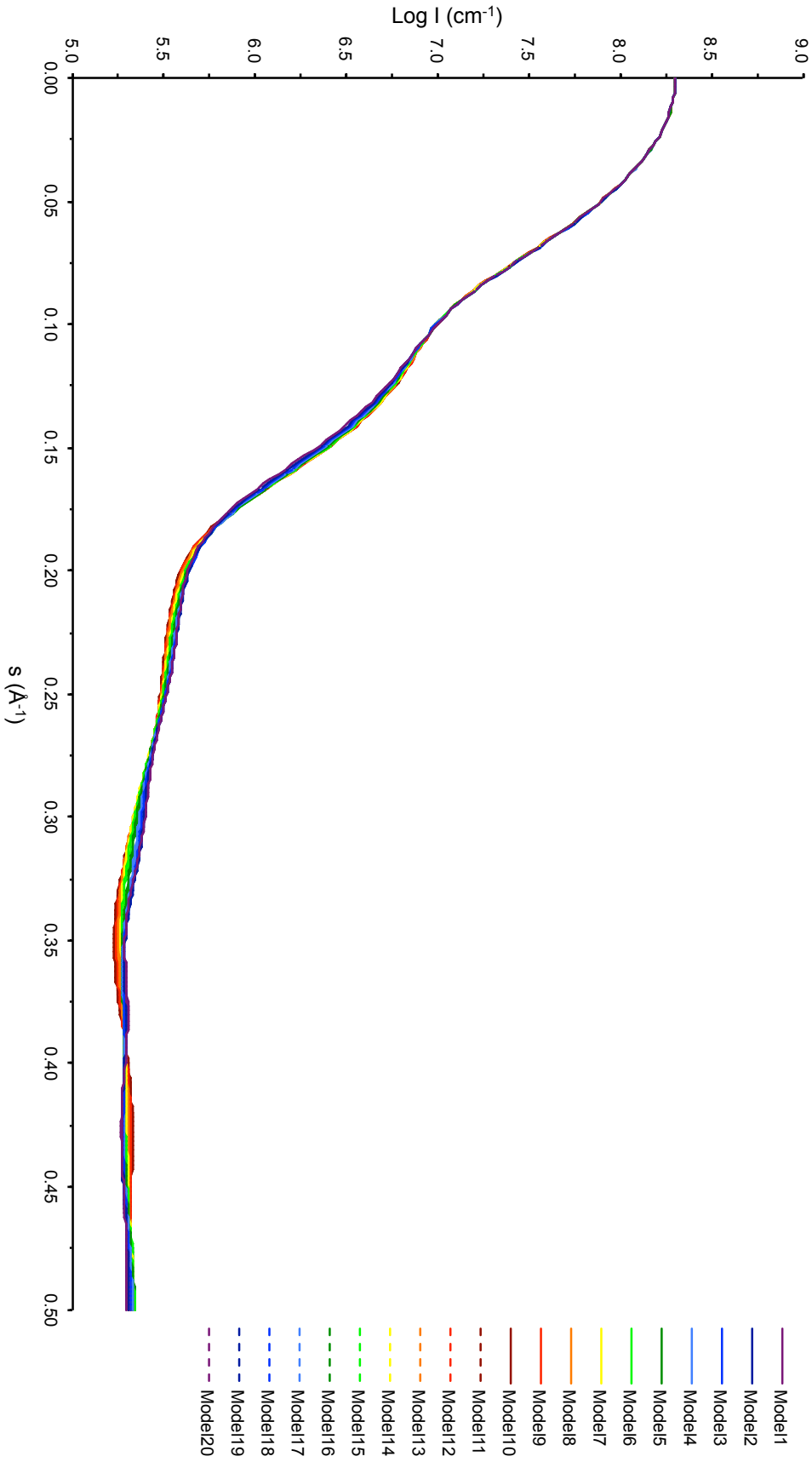


Figure 2.25: Theoretical SAXS scattering for twenty morphed *Nme*IPMS structures, simulating the conversion of conformational extremes in the protein. More symmetrical models are shown in warmer colours, and less symmetrical models in cooler colours.

2.7 Complementary characterisation of *Nme*IPMS

Additional characterisation of *Nme*IPMS has been performed by Michael Hunter;[63] this is summarised here as it contributes to the understanding of *Nme*IPMS. This has revealed that *Nme*IPMS has an activity maximum at a temperature of 40 °C, and is not sensitive to the CoA product inhibition observed in *S. typhimurium* and *S. cerevisiae*.[60,61] This research has also elucidated the metal and α-keto acid specificity of the enzyme.

2.7.1 Metal dependence

As with most forms of α-IPMS that have been characterised to date, *Nme*IPMS displays activation by divalent metals.[41,43,47] Mg^{2+}, Mn^{2+}, Co^{2+} and Ni^{2+} are activating, with Mg^{2+} having the greatest effect. Only Mg^{2+} shows simple saturation kinetics for *Nme*IPMS, whereas other divalent metals have distinct activity maxima, and are less activating at concentrations greater than 200 μM. *Nme*IPMS shows no activation by monovalent cations at any concentration. This is in contrast to *Mtu*IPMS, which demonstrates a strong correlation between monovalent cation concentration and activity, despite the fact that *Mtu*IPMS has no obvious binding site for such a cation.[47]

Analysis of *Nme*IPMS thermal stability in the presence of divalent metal ions shows that 5 mM Mn^{2+}, Co^{2+}, Ni^{2+} or Zn^{2+} increases the denaturation temperature of the enzyme by 10–14 °C, whereas Mg^{2+} and Cd^{2+} have no effect. None of these metal ions significantly altered *Nme*IPMS denaturation temperature at lower concentrations (250 μM).

2.7.2 Substrate specificity

*Nme*IPMS exhibits a similar α-keto acid specificity to *Mtu*IPMS. α-KB and α-ketovalerate have similar turnover rate to the natural substrate (α-KIV) in *Nme*IPMS, but lower affinity, whereas pyruvate has both lower affinity and rate. α-Keto-β-methylvalerate is a weak substrate, whereas glyoxalate and oxaloacetate are weak inhibitors. In *Mtu*IPMS, no activity is observed for α-keto-β-methylvalerate or glyoxalate.[46] *Nme*IPMS, like *Mtu*IPMS,[46] is competitively inhibited by (*S*)-α-hydroxyisovalerate.

2.8 Summary of findings

*Nme*IPMS was successfully cloned, expressed and purified by metal-affinity and size-exclusion chromatography. A CD spectrum indicates this protein is properly folded, and MS shows it is of the appropriate size. This protein exhibits similar properties to other characterised α-IPMS orthologues.

*Nme*IPMS is a dimer in solution in the presence and absence of leucine, as is *Mtu*-IPMS[46] and one isozyme of α-IPMS from *A. thaliana*. Other α-IPMSs have been observed as trimers and tetramers.[41,42,65]

The denaturation temperature of *Nme*IPMS is 44.5 ± 0.1 °C, which is low compared to the 37 °C growth environment for the enzyme. This low thermal stability is perhaps due to the absence of metal ions in the experimental conditions, as these have been shown to increase *Nme*IPMS stability significantly.[63] Increased structural stability of *Nme*IPMS is observed in the presence of the allosteric inhibitor L-leucine or the substrate α-KIV. The denaturation temperature of the enzyme increases by 6.7 ± 0.5 °C in the presence of 250 µM α-KIV, and by 7.7 ± 0.6 °C in the presence of 200 µM L-leucine. Increase in enzyme stability was also seen at high concentrations of DL-valine, and this amino acid shows some inhibition of *Nme*IPMS, decreasing enzyme activity by 6 % at a concentration of 1 mM. No increased stability or enzyme inhibition is observed for *Nme*IPMS in the presence of L-isoleucine or glycine, indicating that feedback inhibition and perhaps allosteric binding is highly selective in this enzyme. These results are similar to those for CMS, which shows inhibitor selectivity for isoleucine over leucine.[9]

Michaelis-Menten kinetic parameters for *Nme*IPMS were found to be similar to those for other α-IPMSs (see Table 1.2 in Section 1.2.1), with binding affinities in the micromolar range and a turnover number of 13 s^{-1}. *Nme*IPMS has a pH activity maximum of 8.5, as observed for other bacterial α-IPMSs.[41,53,57] Like *Mtu*IPMS,[46] *Nme*IPMS is able to catalyse the hydrolysis of the substrate AcCoA in the absence of α-KIV, with reduced AcCoA affinity and a dramatically impaired rate.

Again similar to *Mtu*IPMS, *Nme*IPMS is inhibited by leucine in a mixed, non-competitive manner, however *Nme*IPMS does not demonstrate the time-dependency of inhibition observed for the *M. tuberculosis* enzyme.[45] *Nme*IPMS exhibits pH-dependent sensitivity to leucine inhibition, as observed for *Saccharomyces sp.* and *S. typhimurium*,[53,58] and is almost five times more sensitive to 200 µM L-leucine at pH 6.5 than at pH 8.5.

SAXS data show that *Nme*IPMS adopts a similar conformation in solution to that of the enzyme homology model and, by extension, the *Mtu*IPMS crystal structure the model is based upon. Potential domain movement in the asymmetric dimer was investigated by construction of morphed structural models, however theoretical scattering patterns from these showed no significant changes, and thus SAXS could not be used to assess the asymmetry of the experimental sample.

Chapter 3

Leucine Inhibition in *Nme*IPMS

3.1 Overview

*Nme*IPMS is inhibited by the branched-chain amino acid leucine. This chapter describes experiments conducted to understand this inhibition. Several residues of *Nme*IPMS were substituted for this purpose.

Residue Ser429 was substituted for alanine in an attempt to directly disrupt leucine binding, whereas residues Pro431 and Thr491 were replaced by glycine and alanine, respectively, to better understand the subtleties of structure surrounding the leucine binding site. Substitutions of residues Gly482 and Arg336 to proline and alanine were performed to elucidate the mechanism of transfer of inhibitory signal from the allosteric regulatory site to the catalytic $(\beta/\alpha)_8$-barrel.

Enzyme activity and inhibition were investigated for each of these variants.

3.2 Background

3.2.1 Leucine binding in α-IPMS

As the first committed step of leucine biosynthesis, the reaction catalysed by α-IPMS is controlled by feedback inhibition from the end-product of this pathway. The binding of leucine in the regulatory domain of *Mtu*IPMS has been elucidated by crystallographic studies (PDB code 3FIG).[51]

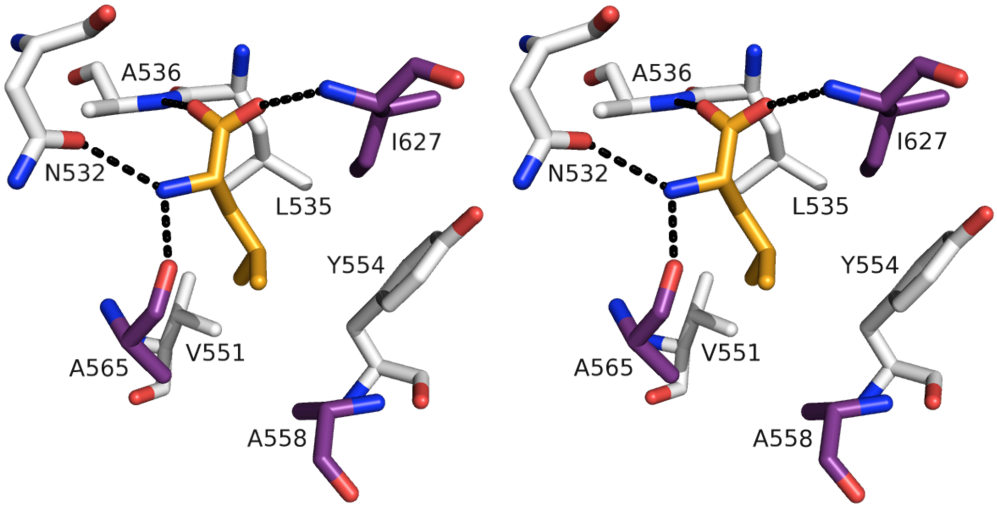


Figure 3.1: Stereo diagram of leucine binding in *Mtu*IPMS structure 3FIG. Each monomer is shown a different colour. Leucine is shown in orange.

In structure 3FIG of *Mtu*IPMS, the amino/carboxylate end of leucine forms hydrogen bonds with four residues. Residue Asn532 (equivalent to Ser429 in *Nme*IPMS) forms a hydrogen bond from its side-chain carbonyl to the amino group of leucine, while Ala536, Ala565′ and Ile627′ (Asp433, Ser463 and Val487 in *Nme*IPMS) form hydrogen bonds to the inhibitor through their main-chain functional groups, as shown in Figure 3.1. Four other residues form the hydrophobic pocket around the methyl groups of leucine — Leu535, Val551, Tyr554 and Ala558′ (Val432, Leu449, Tyr452 and Ala456 in *Nme*IPMS).

Unfortunately some of the secondary structure around the leucine binding site is ill-defined in the *Nme*IPMS homology model, and some side chains of residues that are in the *Mtu*IPMS leucine binding site are oriented away from this site in the *Nme*IPMS model. For this reason, and the fact that the model does not show leucine bound, all *Nme*IPMS residues substituted in this chapter were selected based on the

*Mtu*IPMS leucine-bound structure and a multiple sequence alignment (see Figure 3.7 and Appendix A).

3.2.2 Structural elements in the regulatory domain

The leucine binding site in the *Mtu*IPMS dimer is comprised of residues from the ends of two α-helices (one from each monomer), one β-strand and the flexible loop associated with slow-onset feedback inhibition (Figure 3.2).

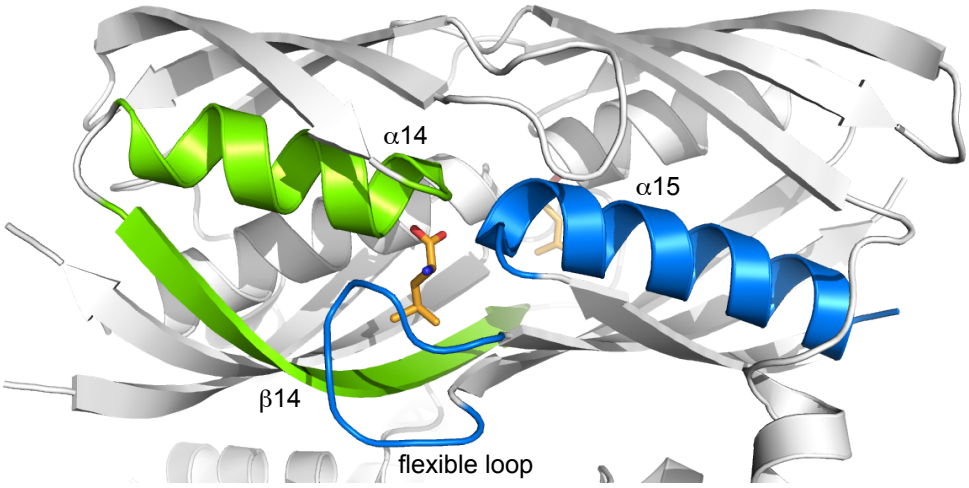


Figure 3.2: Secondary structure in the leucine binding site of *Mtu*IPMS structure 3FIG. Key structural elements are shown in a different colour for each monomer. Leucine is shown in orange.

At the N-terminal start of one of these helices (α14 in Figures 3.2 and 3.7) is a G-x-G-P-[VIL] motif that contributes two residues to the regulatory site; in *Mtu*IPMS the motif residues are Gly531-Asn532-Gly533-Pro534-Leu535 (Figure 3.3), whereas in *Nme*IPMS they are Gly428-Ser429-Gly430-Pro431-Val432. This motif appears to be highly important for enzyme regulation: substitution of either of the glycines (for aspartate) in *S. cerevisiae* α-IPMS results in a leucine-insensitive enzyme,[69] as does substitution of the proline for leucine in *E. coli* α-IPMS.[70]

The first of the *Nme*IPMS substitutions described in this thesis is of Ser429, the variable residue in the G-x-G-P-[VIL] motif and equivalent to *Mtu*IPMS residue Asn532, which forms a hydrogen bond from its carbonyl group to the amino group of leucine (Figures 3.1 and 3.3). The amino acids asparagine or threonine are usually found at this position in α-IPMSs from other organisms. Ser429 in *Nme*IPMS was

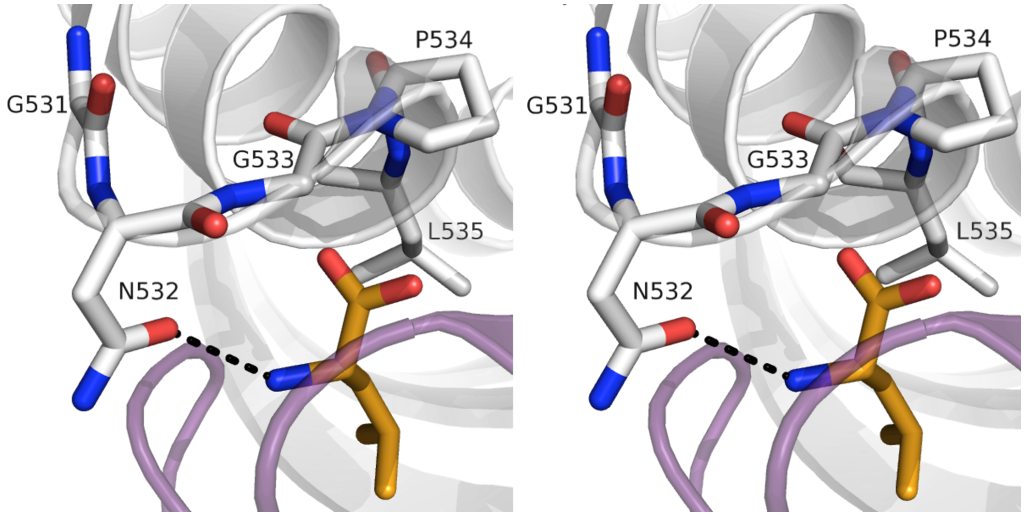


Figure 3.3: Stereo diagram of the G-x-G-P-[VIL] motif in *Mtu*IPMS structure 3FIG. Each monomer is shown in a different colour. Leucine is shown in orange.

substituted for alanine to remove the serine side-chain functionality, and assess the importance of this potential hydrogen bond in *Nme*IPMS sensitivity to leucine.

The second *Nme*IPMS substitution made was of Pro431, the conserved proline in the G-x-G-P-[VIL] motif. Proline is unique amongst the natural amino acids in that its main-chain torsion angle is restricted to about -60°. This dihedral angle constriction is likely important for the backbone structure of the motif loop. Pro431 was substituted for glycine in *Nme*IPMS, which has no dihedral angle restriction. This substitution represents a complete loss of backbone conformation restriction at this position. Although a substitution of the motif proline has already been made in *E. coli* α-IPMS, this previous replacement was with leucine. Leucine has less torsion angle restriction than proline but more than glycine, and also has a bulky hydrophobic side chain that may have additional effects on the enzyme structure. A substitution to glycine frees the main-chain angle restraint of the residue without adding side-chain functionality.

At start of the other leucine-binding-site helix (α15 in Figures 3.2 and 3.7) sits residue Ile627 (Val487 in *Nme*IPMS), which contributes both hydrogen-bonding contacts and hydrophobic contacts to the binding of the inhibitor (Figures 3.1 and 3.4). The main-chain carbonyl of this residue, and of its neighbour Thr628, are stabilised by intra-helix hydrogen bonding to the side-chain hydroxyl and main-chain amino group of Ser631 (Figure 3.4). This serine is highly conserved across α-IPMSs and could ensure correct positioning of Ile627 in the binding pocket. Thr628, on the other hand, is poorly conserved, and equates to residue Leu488 in *Nme*IPMS.

Ser631 is the equivalent of Thr491 in *Nme*IPMS; to determine the importance of this residue in protein function and leucine sensitivity, this threonine was substituted for alanine.

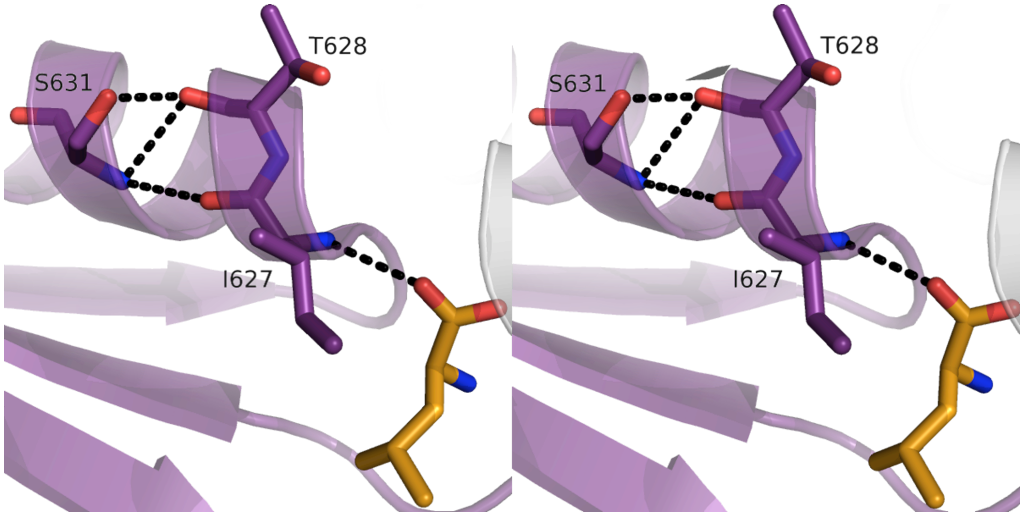


Figure 3.4: Stereo diagram of stabilisation of Ile627 in *Mtu*IPMS structure 3FIG. Each monomer is shown in a different colour. Leucine is shown in orange.

3.2.3 Communication between α-IPMS domains

Multiple sequence alignments reveal that there are two conserved glycines on β-strand β16 in the regulatory domain (Figures 3.5 and 3.7). These glycines are not in direct contact with the leucine binding site, but the strict conservation suggests that they have some functional or structural importance. This β-strand lies at the interface of the regulatory domain and subdomain II, and substitution of one of the glycines for cysteine in *E. coli* α-IPMS results in a leucine-resistant variant.[70]

The regulatory domain to subdomain II interface is composed of hydrophobic interactions from very poorly conserved residues. The flexibility of the conserved glycines could be necessary to allow the surrounding hydrophobic side chains to optimally interact. In *Mtu*IPMS, the lack of a side chain for these glycine residues (Gly620 and Gly622) may also be important for the correct positioning of a conserved alanine (Ala634) that hydrogen bonds to residue Ser631. Through this serine, these glycines may influence the positioning of leucine-binding residue Ile627, as described above. The Gly622 equivalent in *Nme*IPMS (Gly482) was substituted for proline to restrict the main-chain torsion angle and add to the steric bulk at this position.

This substitution probes the structural importance of flexibility and small size in this residue.

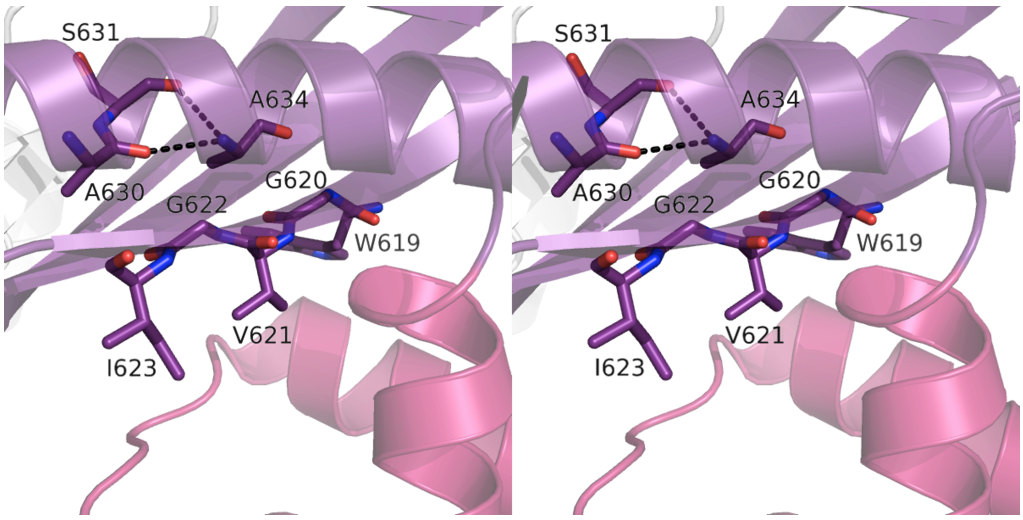


Figure 3.5: Stereo diagram of structural elements at the regulatory to subdomain II interface of *Mtu*IPMS structure 3FIG. One monomer is shown in white, the other is coloured by domain: regulatory domain in purple, subdomain II in pink. The leucine binding site is at the interface of the purple and white regulatory domains, to the left of frame in this diagram.

The final substitution made in *Nme*IPMS was designed to probe potential communication between subdomain II and the catalytic domain. This interface was explored as it seems the most direct pathway for leucine-mediated inhibition to affect the active site.

In *Mtu*IPMS, Lys434 is at the start of subdomain II, and may interact with Asp88, Arg94 or Gln357 in the catalytic $(\beta/\alpha)_8$-barrel, although its side chain is undefined in existing crystal structures (Figure 3.6). Of all the residues in *Mtu*IPMS subdomain II that are positioned to interact with the catalytic domain (including Tyr439 and Glu474), it is the most conserved, appearing as lysine or arginine across all α-IPMS sequences. This residue is an arginine in *Nme*IPMS (Arg336). This was substituted for alanine to examine the importance of the long, basic side chain at this interface.

The interface of the catalytic domain with subdomain II in *Mtu*IPMS is investigated in Chapter 5.

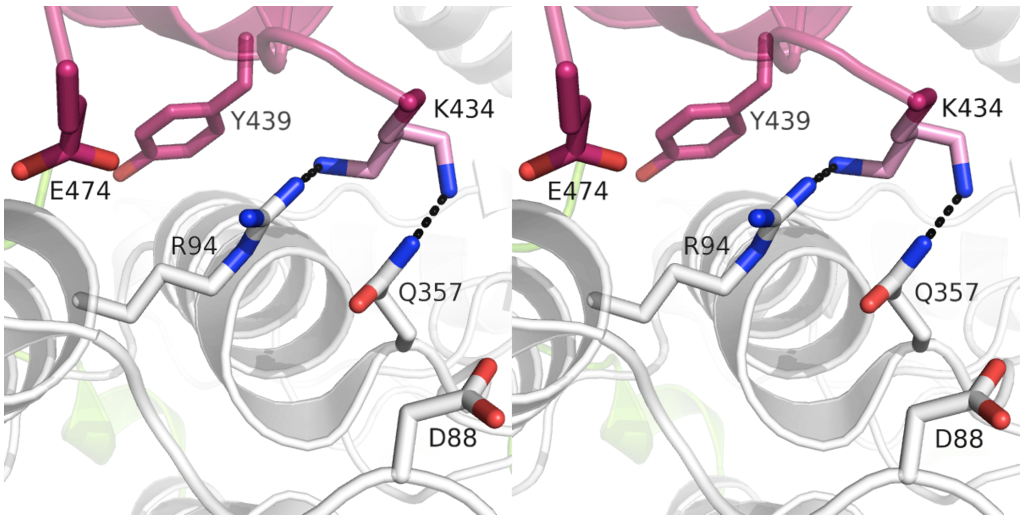


Figure 3.6: Stereo diagram of the subdomain II to catalytic domain interface in *Mtu*IPMS structure 3FIG. The side chain of Lys434 (undefined in the crystal structure) has been built in two possible rotomers. One monomer is shown in white, the other is coloured by domain: subdomain II in pink and the catalytic domain in green.

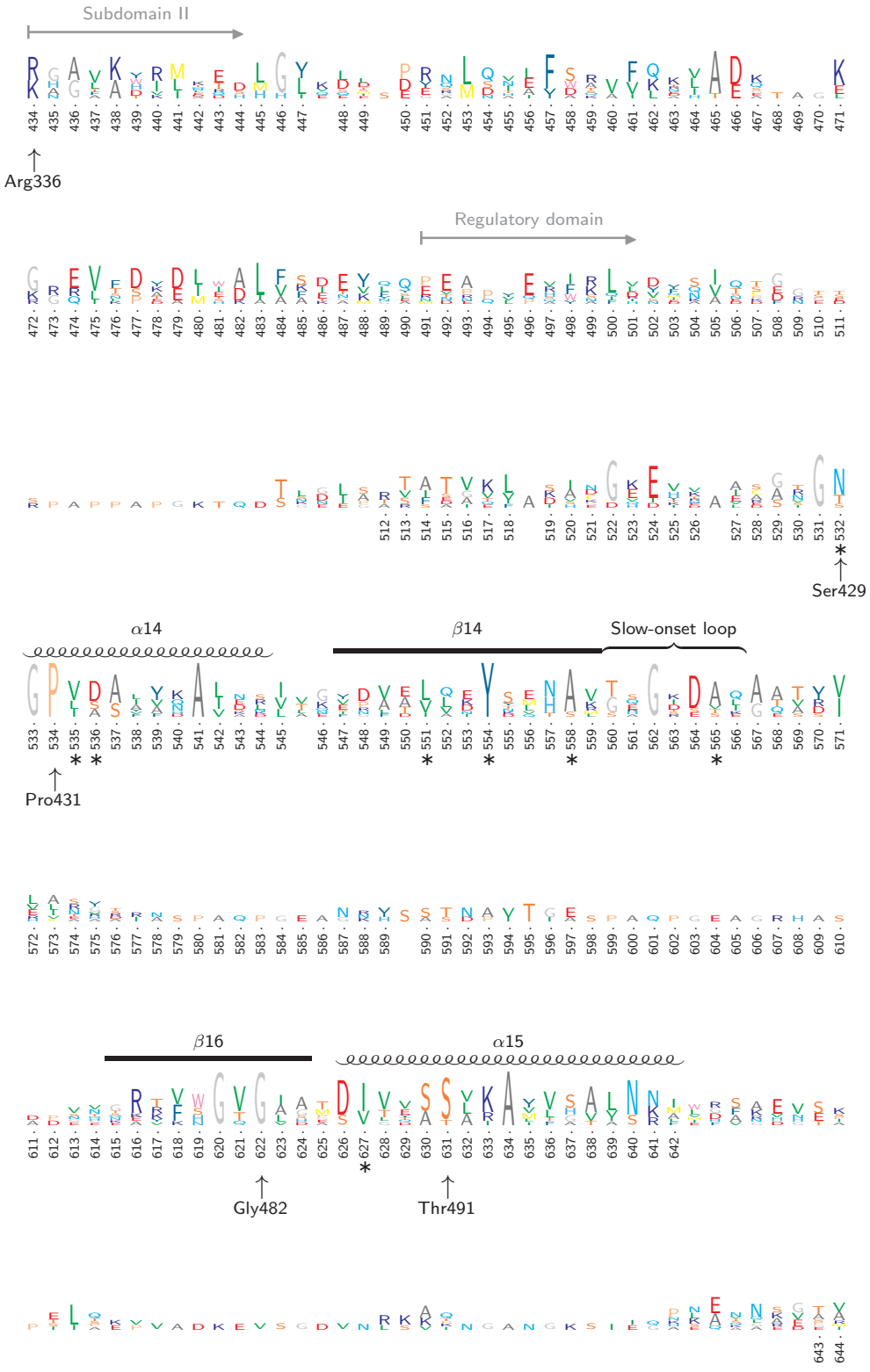


Figure 3.7: Sequence logo of subdomain II and the regulatory domain of α-IPMS. Bottom numbering is for *Mtu*IPMS, with asterisks representing residues involved in leucine binding, and illustrations showing key secondary structure elements in the regulatory domain. *Nme*IPMS residues substituted in this chapter are indicated. Sequences used in construction of this logo are shown in Appendix A.

3.3 Preparation of *Nme*IPMS variants

3.3.1 Cloning

Amino acid substitutions S429A, P431G, G482P, T491A and R336A were generated by introducing the appropriate base-pair changes into the plasmid encoding wild-type *Nme*IPMS.

The mutant plasmids were transformed into *E. coli* XL1 Blue cells, sequenced, and finally transformed into *E. coli* BL21(DE3)Star cells as for the wild-type plasmid.

3.3.2 Expression and purification

Proteins were expressed and purified as for wild-type protein. Variants S429A, P431G, T491A and R336A appeared pure and of a comparable size to wild-type protein by SDS-PAGE.

Variant G482P appeared as a series of bands in SDS-PAGE analysis, even after purification by SEC, and appears to suffer from proteolysis during expression. This proteolysis could be in response to misfolding of the enzyme, which could be due to the structural changes (such as change in main-chain torsion angle or steric bulk) incurred by the glycine to proline substitution. This protein was assessed for activity (see Section 3.5) but was not characterised further than this.

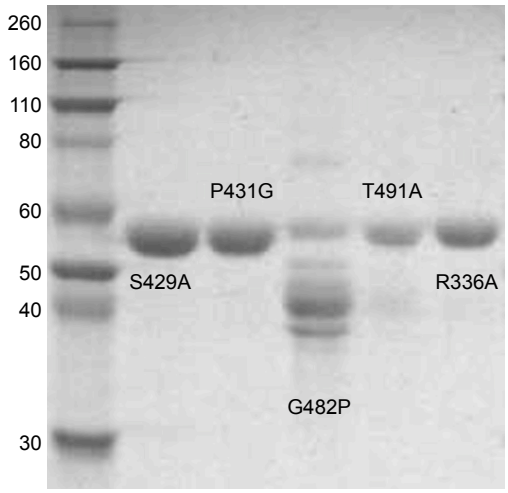


Figure 3.8: SDS-PAGE of *Nme*IPMS variants. Marker weights indicated in kDa.

3.4 Physical characterisation

3.4.1 Presence of secondary structure

CD analysis shows that variants S429A, P431G, T491A and R336A all adopt similar folds to that of wild-type protein (Figure 3.9). G482P was not assessed by CD due to the degradation observed by SDS-PAGE analysis.

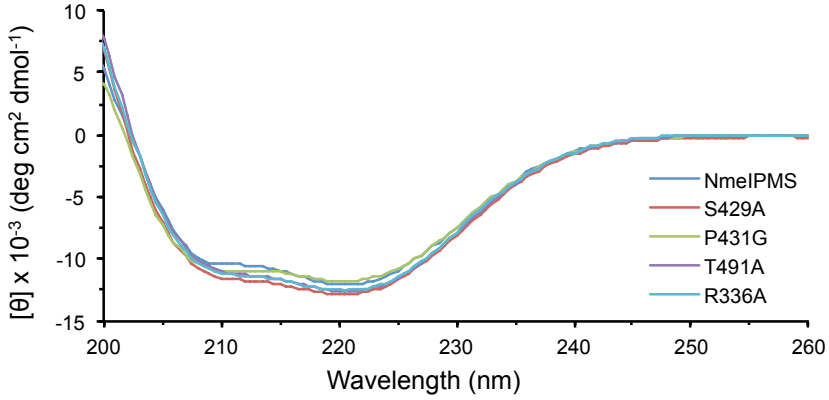


Figure 3.9: Circular dichroism spectra for *Nme*IPMS variants, compared to spectrum of wild-type protein

3.4.2 Molecular mass

All un-proteolysed variants appear to be of an appropriate mass as measured by MS.

	Calculated Mass (Da)	Experimental Mass (Da)
S429A	56011	56010
P431G	55987	55986
T491A	55997	55998
R336A	55942	55941

Table 3.1: Calculated and experimental masses of *Nme*IPMS variants

3.4.3 Thermal stability

All amino-acid substituted variants of *Nme*IPMS showed similar thermal stability to the wild-type protein, as determined by DSF (Table 3.2). The largest change in stability was observed for variant P431G, which showed a 3.9 °C decrease in denaturation temperature when compared to wild-type *Nme*IPMS. It appears that the proline to glycine substitution effects the flexibility of the protein as a whole, in addition to any effects it may have at the site of subsitution.

	T_m (°C)
*Nme*IPMS	44.5 ± 0.1
S429A	45.9 ± 0.3
P431G	40.6 ± 0.7
T491A	45.9 ± 0.1
R336A	42.7 ± 0.2

Table 3.2: Denaturation temperatures (T_m) of *Nme*IPMS variants in 25 mM BTP buffer (pH 7.0), compared to that of wild-type protein.

3.5 Kinetic characterisation

3.5.1 Michaelis-Menten kinetics

Michaelis constants and comparative inhibition data were found for the variants S429A, P431G, T491A and R336A. These variants are all catalytically active with similar turnover numbers to wild-type enzyme. T491A and R336A exhibit slightly decreased maximum rates of catalysis, whereas S429A shows some increase in catalytic activity.

The substitutions present in the *Nme*IPMS variants have minimal effect on α-KIV affinity. In S429A and P431G this affinity is unchanged, whereas in the other two variants a slight decrease in affinity is observed. The most noticeable change between wild-type *Nme*IPMS and these variants is in the reduction of AcCoA affinity. This reduction is largest in P431G and R336A, highlighting the importance of these residues, as discussed below.

G482P was not fully characterised due to the multiple breakdown products seen by SDS-PAGE. Kinetic testing shows it has a turnover rate of $1.6 \pm 0.05\,s^{-1}$ at substrate concentrations of 500 μM, which are saturating concentrations for wild-type protein. Some functional enzyme is thus still present in solution.

	K_m^{app} (μM)			k_{cat}/K_m ($mM^{-1}s^{-1}$)	
	α-KIV	AcCoA	k_{cat} (s^{-1})	α-KIV	AcCoA
*Nme*IPMS	30 ± 2	35 ± 3	12.8 ± 0.3	430 ± 40	370 ± 40
S429A	33 ± 3	45 ± 4	20.6 ± 0.4	640 ± 70	470 ± 50
P431G	28 ± 3	114 ± 18	11.6 ± 0.3	430 ± 60	110 ± 20
T491A	64 ± 7	60 ± 7	7.4 ± 0.2	110 ± 20	120 ± 20
R336A	48 ± 5	800 ± 190	7.0 ± 0.7	150 ± 30	9 ± 3

Table 3.3: Kinetic data for wild-type and variant *Nme*IPMS

The variant P431G has a three-fold higher K_m^{app} for AcCoA than wild-type. This substitution removes the main-chain torsion constriction of the proline residue, potentially affecting the structure of the leucine binding site. As proline is often a key structural element due to its rigid backbone chirality, it is tempting to hypothesise that disruption of structure in the regulatory domain has a flow-on effect through subdomain II to subdomain I, which contributes to the AcCoA binding

pocket in the active site. Any changes to protein structure are likely subtle, in light of the similarity in other kinetic parameters and CD spectrum to those of the wild-type.

The R336A substitution decreases affinity of the enzyme for AcCoA 23-fold, despite being located at least 20 Å from the substrate binding site. This arginine residue may interact with a number of residues in C-terminal loops of the catalytic $(\beta/\alpha)_8$-barrel, some of which are adjacent to the AcCoA binding site. Thus, the interactions between Arg336 and the catalytic domain could act to appropriately configure the active site for AcCoA binding.

3.5.2 Allosteric inhibition

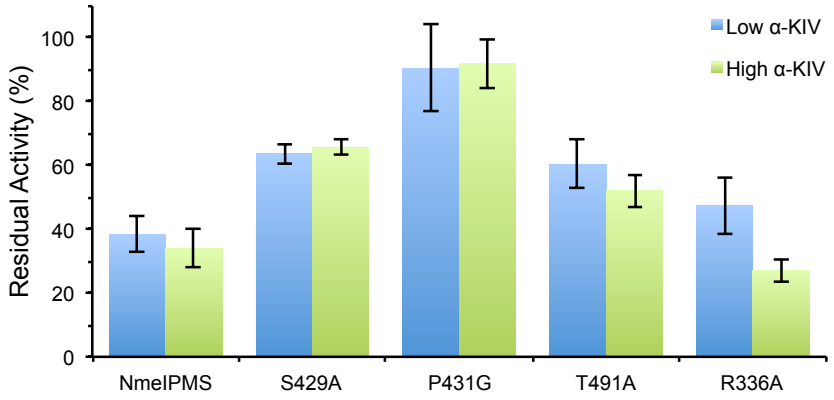
Inhibition by leucine seems relatively unchanged in R336A compared with wild-type *Nme*IPMS (Figure 3.10), but variants S429A and T491A show large desensitisation to leucine, and variant P431G demonstrates almost no sensitivity to this ligand at all.

The effect of substrate concentration on leucine sensitivity in these *Nme*IPMS variants was investigated, and all variants show similar trends to the wild-type in this regard. The reduction in leucine sensitivity observed for S429A and T491A confirms the importance of residues Ser429 and Thr491 in *Nme*IPMS leucine binding. In fact, the loss of leucine sensitivity is very similar for both these variants, despite the fact that Ser429 is likely directly involved in leucine binding, whereas Thr491 plays a secondary role in stabilising a binding residue. This highlights the importance of the residues surrounding the leucine binding pocket in correctly positioning the binding residues themselves.

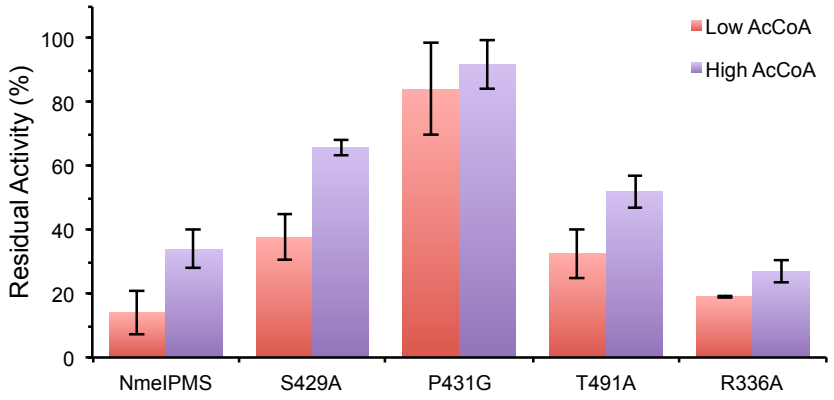
Almost no inhibition is detectable in variant P431G, indicating that the binding-motif proline is critical in maintaining the structure of the inhibitor binding pocket. The flexibility of the P431G variant is compounded by the two other existing glycines in the G-x-G-P-[VIL] motif, potentially allowing a great deal of conformational flexibility in this region.

Pro431 is very near residues Ser429, Val432 and Asp433 (Asn532, Leu535 and Ala536 in *Mtu*IPMS and in Figure 3.1), all of which are part of the leucine binding pocket; with increased flexibility in the motif loop it is likely that at least some of these residues are no longer constrained into a competent conformation for leucine binding. The loss of a serine side chain at position 429 has already been shown above to

decrease leucine sensitivity, but if more than one of the nearby residues are displaced then an entire corner of the leucine binding pocket is lost. This would likely affect leucine binding in the regulatory site.



(a) Inhibition at 500 µM (S429A and T491A) or 1mM (P431G and R336A) AcCoA and varying α-KIV concentrations — 50 and 500 µM



(b) Inhibition at 500 µM α-KIV and varying AcCoA concentrations — 50 and 500 µM for wild-type, S429A and T491A; 100 µM and 1 mM for P431G and R336A

Figure 3.10: Enzyme activity of *Nme*IPMS, S429A, P431G, T491A and R336A in the presence of 1 mM L-leucine. Residual activity is calculated as a percentage of full activity (in the absence of L-leucine) for each variant.

The similarity between inhibition seen for R336A and wild-type *Nme*IPMS indicates that residue Arg336 is not required for the transmission of inhibitory signal from the leucine binding site to the active site. This is despite the fact that the interactions of this arginine side chain at the monomer interface are clearly required for wild-type AcCoA binding affinity.

Arg336 in *Nme*IPMS is equivalent to Lys434 in *Mtu*IPMS. Interestingly, a pep-

tide adjacent to the Lys434-interacting loop in the *M. tuberculosis* enzyme shows changes in molecular dynamics upon leucine binding, as demonstrated by HDX (Figure 3.11).[49] This technique has indicated that residues 78–87 in the AcCoA binding site incorporate fewer deuterium atoms into their backbone-amide hydrogens in the presence of leucine, indicating increased protection from solvent in the presence of the allosteric inhibitor. This peptide contributes two key residues to the active site. It is possible that the R336A substitution in *Nme*IPMS affects AcCoA affinity through this peptide, in a similar mechanism to that of leucine-mediated inhibition.

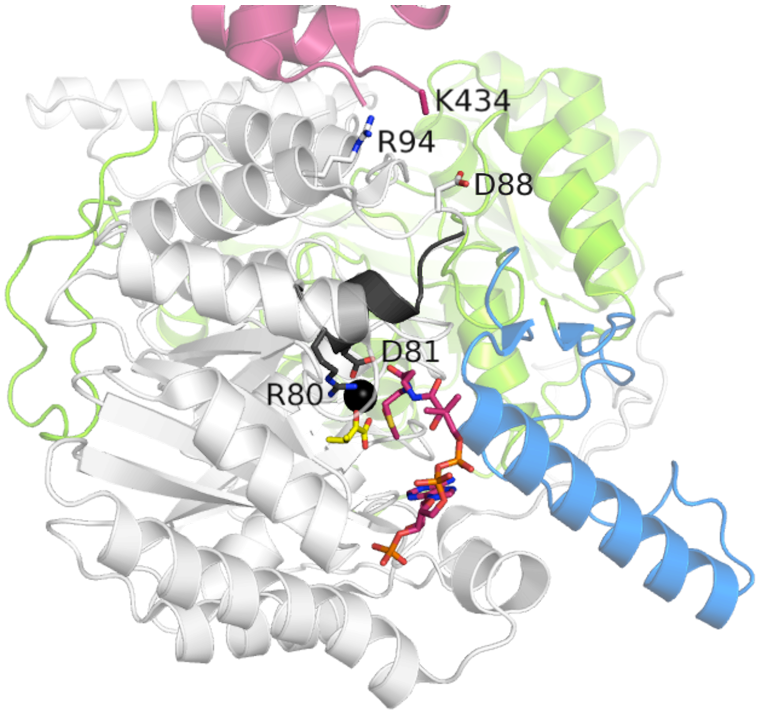


Figure 3.11: Potential mode of Lys434 influence on *Mtu*IPMS AcCoA affinity. One monomer of *Mtu*IPMS structure 3FIG is shown in white, the other is coloured by domain: catalytic domain in green, subdomain I in blue and subdomain II in pink. Zn^{2+} is shown in black, α-KIV in yellow and modelled AcCoA in pink. Residues 78–87 are highlighted in black and the key residues labelled.

3.6 Summary of findings

*Nme*IPMS variants S429A, P431G, T491A and R336A were successfully cloned, expressed and purified as for wild-type protein. These four variants were properly folded and of an appropriate size, with similar denaturation temperatures to the wild-type protein. One further variant, G482P, was cloned and expressed, but appears to suffer from proteolysis and was not fully characterised. This variant possesses some residual activity, but significantly reduced from the activity of wild-type protein.

Substitution of residue Ser429 for alanine results in a variant with similar K_{m} values to wild-type protein and a slightly increased turnover number. This variant has almost twice the residual activity of the wild-type protein in the presence of leucine, indicative of the likely leucine-binding role of this serine in the *Nme*IPMS regulatory domain.

Variant P431G substitutes the conserved proline for glycine in the G-x-G-P-[VIL] motif of the leucine binding site. This variant has unchanged α-KIV affinity and turnover number from the wild-type protein, but shows a 3-fold decrease in affinity for the substrate AcCoA and an almost complete lack of inhibition in the presence of leucine. Substitution of Pro431 for glycine may increase the flexibility in the regulatory domain of the protein — such disruption of structure could then affect the structure and dynamics of the adjacent subdomain II, and through this potentially affect subdomain I and the catalytic $(\beta/\alpha)_8$-barrel. As subdomain I and the $(\beta/\alpha)_8$-barrel form the AcCoA binding site in the protein, disruption to flexibility in this region could represent the mechanism of change in affinity for this substrate. Increased flexibility around Pro431 could also have more localised effects in the regulatory domain, by potentially displacing several leucine-binding residues from their wild-type positions. This displacement could affect leucine binding, which would account for the loss of leucine sensitivity. An increase in protein flexibility is consistent with the decreased thermal stability observed for this variant when compared to wild-type protein.

Similar enzyme parameters to those of variant S429A were found for variant T491A. Thr491 is likely to stabilise leucine-binding residue Val487 in *Nme*IPMS, and is highly conserved as a serine in a multiple sequence alignment of α-IPMS (Figure 3.7 and Appendix A). Substitution of this residue for alanine results in increased K_{m} values for both substrates, and a decreased turnover number compared to wild-type protein. Leucine-sensitivity in this variant is similar to that observed for S429A, both

demonstrating decreased inhibitor sensitivity compared to wild-type protein.

Lastly, interactions at the interface of the catalytic domain with subdomain II in *Nme*IPMS were investigated by substitution of Arg336 for alanine. This variant has slightly lower α-KIV affinity and turnover number than the wild-type protein, and a 23-fold decrease in AcCoA affinity. Unlike variant P431G, in which decreased AcCoA affinity was coupled with loss of leucine-mediated inhibition, R336A shows similar leucine sensitivity to that of wild-type protein. It is possible that the decrease in substrate affinity is mediated through the residues in the C-terminal loops of the $(\beta/\alpha)_8$-barrel, several of which may interact with Arg336 and which are adjacent to the AcCoA binding site.